

# Space-Time Encoded MISO Broadcast Channel with Outdated CSIT: An Error Rate and Diversity Performance Analysis

Bruno Clerckx and David Gesbert

**Abstract**—Studies of the MISO Broadcast Channel (BC) with delayed Channel State Information at the Transmitter (CSIT) have so far focused on the sum-rate and Degrees-of-Freedom (DoF) region analysis. In this paper, we investigate for the first time the error rate performance at finite SNR and the diversity-multiplexing tradeoff (DMT) at infinite SNR of a space-time encoded transmission over a two-user MISO BC with delayed CSIT. We consider the so-called MAT protocol obtained by Maddah-Ali and Tse, which was shown to provide 33% DoF enhancement over TDMA. While the asymptotic DMT analysis shows that MAT is always preferable to TDMA, the Pairwise Error Probability analysis at finite SNR shows that MAT is in fact not always a better alternative to TDMA. Benefits can be obtained over TDMA only at very high rate or once concatenated with a full-rate full-diversity space-time code. The analysis is also extended to spatially correlated channels and the influence of transmit correlation matrices and user pairing strategies on the performance are discussed. Relying on statistical CSIT, signal constellations are further optimized to improve the error rate performance of MAT and make it insensitive to user orthogonality. Finally, other transmission strategies relying on delayed CSIT are discussed.

**Index Terms**—MISO Broadcast Channel, delayed CSIT, Diversity-Multiplexing Tradeoff, performance analysis, multiuser communications, space-time coding, user pairing

## I. INTRODUCTION

THE performance of multiuser MISO/MIMO highly depends on the accuracy of Channel State Information at the Transmitter (CSIT) [1], [2]. Obtaining accurate CSIT is however challenging in practice as the feedback is subject to various impairments including delay. There has been recent progress on understanding how perfect but outdated CSIT can be beneficial to boost the Degrees-of-Freedom (DoF) of MISO Broadcast Channels [3]. It was shown that in a two-user MISO BC with outdated CSIT, a sum DoF of  $4/3$  can be achieved using a transmission strategy that exploits the knowledge of the outdated CSIT to reconstruct and transmit the interference overheard by co-scheduled receivers. This leads to a 33% DoF enhancement compared to conventional TDMA approach. Throughout the paper, we will denote such transmission strategy as MAT. In [4], then in [5], [6], authors extended the work by looking at a scenario with both imperfect instantaneous

CSIT and perfect delayed CSIT where the idea of imperfect current CSIT lies in the fact that prediction can be applied to delayed CSIT. An alternative transmission strategy, denoted as Alternative MAT in this paper, was suggested to also achieve a sum DoF of  $4/3$  and used as a building block of a more general strategy suitable for time correlated channels. [7] generalized the findings in [3] by giving an optimal DoF region for an alternative CSIT setting in a two-user, where the CSIT of each user can be perfect, delayed or absent. In [8], the results of [5], [6] were extended by considering different qualities of instantaneous CSIT of the two users. Further analysis on the impact of imperfect outdated CSIT on the DoF region was made in [9]. Recently, DoF results found in the two-user time-correlated MISO BC with delayed CSIT have been extended to the MIMO and interference channel cases in [10]. So far, all cited works exclusively focused on DoF analysis at high SNR. In [11], a sum-rate criterion is considered and the design of precoders to enhance the performance of MAT at finite SNR is addressed. In [12], [13], a similar criterion is chosen but the precoders are designed based on statistical CSIT. In [14], authors depart from classical DoF and rate analysis and also investigate diversity performance of MAT strategy. An outer-bound on the asymptotic diversity-multiplexing tradeoff (DMT) at infinite SNR achieved by MAT is derived and a novel scheme is proposed to achieve both full DoF as well as full diversity.

Surprisingly, several transmission strategies have been derived to cope with delayed CSIT in the cited references and much is known about their DoF performance but the diversity performance (including the asymptotic DMT) has been overlooked so far. In particular, the error rate performance of corresponding techniques at finite SNR in MISO BC with outdated CSIT has never been addressed so far. The objective of this paper is to fill this gap by contrasting the asymptotic DMT behavior and the finite SNR error rate performance and identify whether gains over the scenario where delayed CSIT is simply ignored (e.g. TDMA) are always observed. Specifically, the contributions of the paper are listed as follows:

- For a two-user MISO BC, the average Pairwise Error Probability (PEP) over spatially correlated (at the transmitter) Rayleigh fading channels is derived for a general space-time encoded MAT strategy. The PEP behaviour is discussed as a function of the SNR and the transmit correlation matrices of both users and contrasted with conventional behaviours of multi-user MIMO with perfect

Bruno Clerckx is with Imperial College London, London SW7 2AZ, United Kingdom and also with the School of Electrical Engineering, Korea University, Seoul 136-701, Korea (email: b.clerckx@imperial.ac.uk). David Gesbert is with EURECOM, Sophia-Antipolis, France (email: david.gesbert@eurecom.fr). This work was supported in part by the Seventh Framework Programme for Research of the European Commission under grant number HARP-318489.

CSIT and point-to-point MIMO.

- The exact diversity-multiplexing tradeoff (DMT) and the one achieved by various space-time encoded MAT in i.i.d. Rayleigh fading channels are identified and a suitable space-time code design criterion is proposed. Interestingly, the exact DMT of MAT is shown to match the upper bound found in [14].
- Relying on the PEP analysis, a suitable user pairing strategy is discussed. Pairing statistically orthogonal users with similar magnitudes of the transmit correlation coefficients is shown to be a good strategy. With such a strategy, transmit correlation is shown to have a minor impact on the error rate performance of MAT over a large range of SNR.
- If transmit spatial correlations are known at the transmitter, following a previous work by the authors [17], signal constellations are further optimized to improve the error rate performance of MAT in spatially correlated channels. The performance with such constellations are shown to be insensitive to user orthogonality.
- The error rate and DMT performance analysis is finally extended to other strategies relying on delayed CSIT, namely the Alternative MAT proposed in [5]. Despite the fact that both MAT and Alternative MAT have the same multiplexing gain (DoF), Alternative MAT is shown to incur lower diversity and coding gains than the original MAT.

Overall, the primary takeaway observation from this work is that in i.i.d. Rayleigh fading channels, MAT always outperforms TDMA (that would ignore the delayed CSIT) from an asymptotic DMT perspective but is not always a better alternative to TDMA from an error rate perspective at finite SNR. Benefits are obtained over TDMA only at very high rate (no gains have been identified at a per-user rate of 3 bit/s/Hz or below) or once concatenated with a full-rate full-diversity space-time code. Alternative MAT, while being superior to TDMA from an asymptotic DMT perspective, is on the other hand outperformed by TDMA over a wide SNR and rate range when it comes to error probability performance. In spatially correlated channels, MAT with a suitable user pairing strategy is shown to be less sensitive to transmit spatial correlation than TDMA.

The paper is organized as follows. Section II details the system model. Section III details the error rate performance in i.i.d. Rayleigh fading channels and discusses the diversity-multiplexing tradeoff and code design criterion. Section IV extends the analysis to spatially correlated channels and derives user pairing strategies and optimized signal constellations. Section V extends the discussions to alternative schemes and Section VI confirms through simulations the observations made from the analytical derivations. Section VII concludes the work.

The following notations are used throughout the paper. Bold lower case and upper case letters stand for vectors and matrices respectively whereas a symbol not in bold font represents a scalar.  $(\cdot)^T$  and  $(\cdot)^H$  represent the transpose and conjugate transpose of a matrix or vector respectively. Operators  $\det(\cdot)$  and  $\text{Tr}\{\cdot\}$  refer to the determinant and trace of a matrix, res-

spectively.  $\mathcal{E}\{\cdot\}$  refers to the expectation operator. Logarithm log is taken in base  $e$ , unless otherwise specified.  $\text{diag}\{\mathbf{a}\}$  refers to the diagonal matrix whose elements are given by the entries of  $\mathbf{a}$ . Finally,  $f(\rho) \doteq g(\rho)$  indicates  $\lim_{\rho \rightarrow \infty} \frac{f(\rho)}{g(\rho)} = 1$ .  $\leq$  is defined similarly.

## II. SYSTEM MODEL

Let us assume a two-user two transmit antenna MISO BC with delayed CSIT. The transmission occurs over three coherence times. Each coherence time is made of  $T$  time slots over which the channel is constant. We denote the channel vector of user 1 on coherence time  $k$  as  $\mathbf{h}_k = [h_{k,1} \ h_{k,2}]$  where any entry  $h_{k,m}$  refers to the channel coefficient from transmit antenna  $m$ . Similarly,  $\mathbf{g}_k = [g_{k,1} \ g_{k,2}]$  and  $g_{k,m}$  are defined for user 2.

Denoting the transmit signal on time slot  $t$  of coherence time  $k$  as  $\mathbf{x}_{k,t}$ , the received signals at user 1 and 2, respectively denoted as  $y_{k,t}$  and  $z_{k,t}$ , write as

$$y_{k,t} = \mathbf{h}_k \mathbf{x}_{k,t} + n_{k,t}, \quad (1)$$

$$z_{k,t} = \mathbf{g}_k \mathbf{x}_{k,t} + w_{k,t}, \quad (2)$$

where  $n_{k,t} \sim \mathcal{CN}(0, 1)$  and  $w_{k,t} \sim \mathcal{CN}(0, 1)$  are AWGN. We consider a long-term power constraint  $\mathcal{E}\{\mathbf{x}_{k,t}^H \mathbf{x}_{k,t}\} \leq \rho_{lt}$ .

The channel coefficients are modeled as identically distributed circularly symmetric complex Gaussian variables but can be either independent as in Section III or spatially correlated as in Section IV. They are assumed constant within a coherence time and change independently from one coherence time to the next one. The CSI is assumed to be available at the transmitter only at the next coherence time. For simplicity, the normalization  $\mathcal{E}\{|h_{k,m}|^2\} = \mathcal{E}\{|g_{k,m}|^2\} = 1$  is made. We define the transmit covariance matrices  $\mathbf{R}_{t,1}$  and  $\mathbf{R}_{t,2}$  for user 1 and 2 respectively as

$$\begin{aligned} \mathbf{R}_{t,1} &= \mathcal{E}\{\mathbf{h}_k^H \mathbf{h}_k\} = \begin{bmatrix} 1 & t_1^* \\ t_1 & 1 \end{bmatrix}, \quad \forall k \\ \mathbf{R}_{t,2} &= \mathcal{E}\{\mathbf{g}_k^H \mathbf{g}_k\} = \begin{bmatrix} 1 & t_2^* \\ t_2 & 1 \end{bmatrix}, \quad \forall k. \end{aligned} \quad (3)$$

The quantities  $t_1$  and  $t_2$  are the transmit correlation coefficients and can be expressed in terms of their magnitudes and phases as  $t_1 = |t_1| e^{j\varphi_1}$  and  $t_2 = |t_2| e^{j\varphi_2}$ . Whenever we assume i.i.d. circularly symmetric complex Gaussian variables (denoted in short as i.i.d. Rayleigh fading in the sequel),  $\mathbf{R}_{t,1}$  and  $\mathbf{R}_{t,2}$  are identity matrices.

### A. From MAT to Space-Time Encoded MAT

We consider two independent codewords,  $\mathbf{C} = [\mathbf{c}_1, \dots, \mathbf{c}_T]$  and  $\mathbf{C}' = [\mathbf{c}'_1, \dots, \mathbf{c}'_T]$  respectively intended for user 1 and user 2. Their sizes are  $2 \times T$ , therefore spanning the two transmit antennas and  $T$  time slots. The codewords are normalized such that  $\mathcal{E}\{\text{Tr}\{\mathbf{C}\mathbf{C}^H\}\} = \mathcal{E}\{\text{Tr}\{\mathbf{C}'\mathbf{C}'^H\}\} = T$ . The aim is to transmit codeword  $\mathbf{C}$  to user 1 and codeword  $\mathbf{C}'$  to user 2 over  $3T$  time slots using the MAT strategy [3]. The system can therefore be viewed as a space-time encoded transmission over a two-user MISO BC with delayed CSIT. For simplicity, we will look at the performance of the first user only.

In MAT, it is commonly assumed that the codewords  $\mathbf{C}$  and  $\mathbf{C}'$  are chosen as in Spatial Multiplexing (SM), i.e. independent data streams are transmitted from each transmit antenna, because the focus is on DoF maximization. Assuming an uncoded (no FEC) transmission, the time slot index  $t$  can be dropped and the  $2 \times 1$  codewords  $\mathbf{C} = \mathbf{c}$  and  $\mathbf{C}' = \mathbf{c}'$  span only one symbol duration ( $T = 1$ ). As detailed in [3], MAT transmits  $\mathbf{x}_1 = \sqrt{\rho}\mathbf{c}$  in coherence time 1,  $\mathbf{x}_2 = \sqrt{\rho}\mathbf{c}'$  in coherence time 2 and the overheard interference  $\mathbf{x}_3 = \sqrt{\rho}[\mathbf{g}_1\mathbf{c} + \mathbf{h}_2\mathbf{c}' \ 0]^T$  in coherence time 3. A long-term average transmit power (where averaging is also taken over the channel realizations) of  $\rho_{lt} = 4/3\rho$  is consumed and twice as much power is spent on coherence time 3 as in coherence time 1 and 2. In the sequel, we will refer to  $\rho$  as the SNR. It results in the following equivalent system model for user 1

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \sqrt{\rho} \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{0} \\ h_{3,1}\mathbf{g}_1 \end{bmatrix} \mathbf{c} + \sqrt{\rho} \begin{bmatrix} \mathbf{0} \\ \mathbf{h}_2 \\ h_{3,1}\mathbf{h}_2 \end{bmatrix} \mathbf{c}' + \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix}. \quad (4)$$

After further interference elimination,

$$\tilde{\mathbf{y}} = \begin{bmatrix} y_1 \\ y_3 - h_{3,1}y_2 \end{bmatrix} = \sqrt{\rho}\mathbf{H}\mathbf{c} + \begin{bmatrix} n_1 \\ n_3 - h_{3,1}n_2 \end{bmatrix} \quad (5)$$

where

$$\mathbf{H} = \begin{bmatrix} h_{1,1} & h_{1,2} \\ h_{3,1}g_{1,1} & h_{3,1}g_{1,2} \end{bmatrix}. \quad (6)$$

This is an equivalent  $2 \times 2$  MIMO channel and with the instantaneous channel realizations perfectly known at the receive side and shared across users (as  $\mathbf{H}$  is a function of both users' channels), an estimate of the entries of  $\mathbf{c}$  can be obtained using e.g. a Maximum-Likelihood (ML) or linear receiver. This transmission strategy is denoted as SM-encoded MAT in the sequel.

However, MAT as presented in [3] is a framework that is applicable to any number  $T$  of time slots and is not limited to SM-type of codewords. Hence, in the rest of this paper, we will often refer to O-STBC encoded MAT or more generally space-time encoded MAT to stress that the codewords  $\mathbf{C}$  and  $\mathbf{C}'$  are either chosen as O-STBC or from a general space-time code (including SM-encoded MAT). On time slot  $t$ , a space-time encoded MAT consists in transmitting  $\mathbf{x}_{1,t} = \sqrt{\rho}\mathbf{c}_t$  in coherence time 1,  $\mathbf{x}_{2,t} = \sqrt{\rho}\mathbf{c}'_t$  in coherence time 2 and the overheard interference  $\mathbf{x}_{3,t} = \sqrt{\rho}[\mathbf{g}_1\mathbf{c}_t + \mathbf{h}_2\mathbf{c}'_t \ 0]^T$  in coherence time 3. The equivalent system model for user 1 at time instant  $t = 1, \dots, T$  can then be written as

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \end{bmatrix} = \sqrt{\rho} \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{0} \\ h_{3,1}\mathbf{g}_1 \end{bmatrix} \mathbf{c}_t + \sqrt{\rho} \begin{bmatrix} \mathbf{0} \\ \mathbf{h}_2 \\ h_{3,1}\mathbf{h}_2 \end{bmatrix} \mathbf{c}'_t + \begin{bmatrix} n_{1,t} \\ n_{2,t} \\ n_{3,t} \end{bmatrix} \quad (7)$$

and after further interference elimination

$$\tilde{\mathbf{y}}_t = \begin{bmatrix} y_{1,t} \\ y_{3,t} - h_{3,1}y_{2,t} \end{bmatrix} = \sqrt{\rho}\mathbf{H}\mathbf{c}_t + \begin{bmatrix} n_{1,t} \\ n_{3,t} - h_{3,1}n_{2,t} \end{bmatrix}. \quad (8)$$

Users perceive a space-time encoded transmission over an equivalent  $2 \times 2$  MIMO channel. We note the difference in terms of channel matrix  $\mathbf{H}$  in (6) and noise compared to

a classical  $2 \times 2$  point-to-point space-time encoded MIMO system model [18]. In particular, the entries of the channel matrix  $\mathbf{H}$  are not identically distributed and are function of both users' channels. Those differences make the performance analysis (including the error rate, diversity-multiplexing trade-off, impact of spatial correlation) different from conventional point-to-point MIMO Rayleigh fading channels and multi-user MISO with perfect CSIT.

### B. Performance Metrics

Since the encoded transmission is performed over a small number of channel realizations and the transmitter does not have perfect CSIT, outage may occur. Hence outage and error probabilities are valid performance metrics that will be used throughout this paper. The outage probability is used to identify the asymptotic DMT and the error probability is used to characterize the behavior of practical space-time encoded MAT strategies at finite SNR.

It is assumed that ML decoding is performed at receiver 1 (and similarly for receiver 2) in order to estimate  $\mathbf{C}$  from the received signals  $\tilde{\mathbf{y}}_t$ ,  $t = 1, \dots, T$ . With instantaneous channel realizations perfectly known at the receive side and shared across users, the ML decoder at receiver 1 computes an estimate of the transmitted codeword according to

$$\hat{\mathbf{C}} = \arg \min_{\mathbf{C}} \sum_{t=1}^T \left\| \Sigma^{-1/2} (\tilde{\mathbf{y}}_t - \sqrt{\rho}\mathbf{H}\mathbf{c}_t) \right\|^2 \quad (9)$$

where  $\Sigma = \text{diag}\{1, 1 + |h_{3,1}|^2\}$  is the covariance matrix of the noise vector in (8). The minimization in (9) is performed over all possible codewords.

Let us define the following matrix  $\tilde{\mathbf{H}}$

$$\tilde{\mathbf{H}} = \Sigma^{-1/2}\mathbf{H} = \begin{bmatrix} 1 & 0 \\ 0 & X \end{bmatrix} \underbrace{\begin{bmatrix} h_{1,1} & h_{1,2} \\ g_{1,1} & g_{1,2} \end{bmatrix}}_{\mathbf{H}'} \quad (10)$$

where  $X = \frac{h_{3,1}}{\sqrt{1+|h_{3,1}|^2}}$ . When a codeword  $\mathbf{C}$  is transmitted, we are interested in the error probability, called Pairwise Error Probability (PEP), that the ML decoder decodes the codeword  $\mathbf{E}$  instead of  $\mathbf{C}$ . The conditional PEP  $P(\mathbf{C} \rightarrow \mathbf{E} | \tilde{\mathbf{H}})$  with a ML decoder can then be written as

$$P(\mathbf{C} \rightarrow \mathbf{E} | \tilde{\mathbf{H}}) = \mathcal{Q} \left( \sqrt{\frac{\rho}{2}} \left\| \tilde{\mathbf{H}}(\mathbf{C} - \mathbf{E}) \right\|_F \right). \quad (11)$$

Even though the analysis could be conducted based on the exact PEP (by making use of the Craig's formula [1]), for the sake of readability and simplicity, we will make use of the Chernoff bound and therefore upper bound the PEP as

$$P(\mathbf{C} \rightarrow \mathbf{E} | \tilde{\mathbf{H}}) \leq \exp \left( -\frac{\rho}{4} \left\| \tilde{\mathbf{H}}(\mathbf{C} - \mathbf{E}) \right\|_F^2 \right). \quad (12)$$

### III. UNCORRELATED FADING CHANNELS

We discuss spatially uncorrelated and correlated scenarios in two different sections. We start here with the uncorrelated case, i.e.  $\mathbf{R}_{t,1} = \mathbf{R}_{t,2} = \mathbf{I}_2$ . The average PEP for both MAT is obtained by taking the expectation of the conditional PEP



(12) over the channel distribution. The expectation can be computed in two steps: first by taking the expectation over the distribution of  $\mathbf{H}'$  and then over the distribution of  $X$ .

#### A. Error Rate Performance of (encoded) MAT

Following the derivations in Appendix A, we get the upper bound on the average PEP of space-time encoded MAT, as displayed in (13), by taking the expectation over the channel distribution.  $\text{Ei}(x)$  is the exponential integral and  $b_k = 1 + a_k$  ( $k = 1, 2$ ) with  $a_k = \frac{\rho}{4}\lambda_k$  and  $\lambda_k$  the  $k^{\text{th}}$  eigenvalue of the error matrix  $\tilde{\mathbf{E}} = (\mathbf{C} - \mathbf{E})(\mathbf{C} - \mathbf{E})^H$ .

Assuming a full-rank code, i.e.  $\lambda_k > 0$  for  $k = 1, 2$ , the average PEP at sufficiently high SNR can be approximated as

$$P(\mathbf{C} \rightarrow \mathbf{E}) \leq \left(\frac{\rho}{4}\right)^{-3} (\lambda_1 \lambda_2)^{-1} \frac{\log(\lambda_1) - \log(\lambda_2)}{\lambda_1 - \lambda_2}. \quad (14)$$

The maximum achievable diversity gain is 3 and the coding gain is given by  $(\lambda_1 \lambda_2)^{-1} \frac{\log(\lambda_1) - \log(\lambda_2)}{\lambda_1 - \lambda_2}$  where  $\lambda_1 \lambda_2 = \det(\tilde{\mathbf{E}})$ . We note the difference with the classical rank-determinant criterion in space-time code design over i.i.d. Rayleigh fading channels [18]. Intuitively, a maximum diversity gain of 3 (rather than 4) is achieved with a full-rank code because of the presence of  $h_{3,1}$  in both entries of the second row of  $\mathbf{H}$  in (6). This implies that the entries of the second row of (6) do not fade independently and that an error is likely to occur whenever the three channel coefficients  $h_{3,1}$ ,  $h_{1,1}$  and  $h_{1,2}$  are in deep fade.

We particularize the result to the following two cases of Spatial Multiplexing and Orthogonal-Space Time Block Codes (O-STBC).

Let us first assume a SM-encoded MAT where the space-time transmission is operated using Spatial Multiplexing (SM) with independent streams transmitted from each antenna without coding across antennas. Such transmission leads to rank-1 error matrix  $\tilde{\mathbf{E}}$  with a unique non-zero eigenvalue  $\lambda$ . Hence  $\lambda_2 = 0$ ,  $b_2 = 1$ ,  $b_1 = b = 1 + a$  and  $b_1 - b_2 = a$  with  $a = \frac{\rho}{4}\lambda$ . Denoting  $\mathbf{C} - \mathbf{E} = \frac{1}{\sqrt{2}} \begin{bmatrix} c_0 - e_0 & c_1 - e_1 \end{bmatrix}^T$ , we get  $\lambda = \frac{1}{2} \left[ |c_0 - e_0|^2 + |c_1 - e_1|^2 \right]$ . The average PEP (13) simplifies into

$$P(\mathbf{C} \rightarrow \mathbf{E}) \leq \frac{1}{b} \left[ \frac{1}{b} - \frac{a}{b^2} \exp\left(\frac{1}{b}\right) \text{Ei}\left(\frac{-1}{b}\right) \right]. \quad (15)$$

At high SNR,

$$P(\mathbf{C} \rightarrow \mathbf{E}) \leq \frac{\log(a)}{a^2} = \left(\frac{\rho}{4}\right)^{-2} \lambda^{-2} \log\left(\frac{\rho}{4}\lambda\right). \quad (16)$$

Focusing on the worst case PEP (i.e. maximum PEP among all possible pairs of codewords  $\mathbf{C}$  and  $\mathbf{E}$  with  $\mathbf{C} \neq \mathbf{E}$ ), the average error probability at high SNR of SM-encoded MAT can be approximated as

$$P_{SM-MAT} \approx \left(\frac{\rho}{4}\right)^{-2} \lambda_{\min}^{-2} \log\left(\frac{\rho}{4}\lambda_{\min}\right) \quad (17)$$

with  $\lambda_{\min} = \min_{\mathbf{C} \neq \mathbf{E}} \frac{1}{2} \left[ |c_0 - e_0|^2 + |c_1 - e_1|^2 \right] = \frac{1}{2} d_{\min,M}^2$  where  $d_{\min,M}^2$  refers to the squared minimum distance of a constellation with  $2^{\frac{3}{2}R}$  points aiming to achieve

a per-user rate  $R$  (with two symbols transmitted over three channel uses). Hence,

$$P_{SM-MAT} \approx \left(\frac{\rho d_{\min,M}^2}{8}\right)^{-2} \log\left(\frac{\rho d_{\min,M}^2}{8}\right). \quad (18)$$

From (16), the diversity gain writes as  $d = -\frac{\partial \log(P(\mathbf{C} \rightarrow \mathbf{E}))}{\partial \log(a)} = 2 - \frac{1}{\log(a)} \approx 2$ . Hence in the limit of infinite SNR, a diversity gain of 2 is achievable. Strictly speaking, due to the double Rayleigh distribution of some channel coefficients, the slope of the error probability is not as steep as with the classical Rayleigh distributed MIMO/MISO channel, hence leading to a diversity gain that appears slightly lower than 2 at finite SNR. Intuitively, the diversity gain of 2 comes from the fact that the receiver still has access to two independent observations of the transmitted codeword despite the presence of  $h_{3,1}$  in both entries of the second row of (6).

With O-STBC (Alamouti code) [15] encoded MAT,  $\tilde{\mathbf{E}} = \alpha \mathbf{I}_2$ ,  $b_1 = b_2 = b = 1 + a = 1 + \frac{\rho}{4}\alpha$  and the average PEP is written as

$$P(\mathbf{C} \rightarrow \mathbf{E}) \leq \frac{1}{b^2} \left[ \frac{b^3 - b^2 + b}{b^4} + \frac{(1-b)^2}{b^4} \exp\left(\frac{1}{b}\right) \text{Ei}\left(\frac{-1}{b}\right) \right], \quad (19)$$

which leads at high SNR to

$$P(\mathbf{C} \rightarrow \mathbf{E}) \leq \frac{1}{a^2} \left[ \frac{1}{a} + \frac{\log(a)}{a^2} \right] \approx \frac{1}{a^3} = \left(\frac{\rho}{4}\right)^{-3} \alpha^{-3}. \quad (20)$$

A diversity gain of 3 is achieved.

We note that a simple TDMA transmission that ignores the delayed CSIT would achieve a diversity gain of 2 by simply transmitting to each user at a time using O-STBC within a coherence time. In an i.i.d. Rayleigh fading MISO channel with two transmit antennas, focusing on the worst-case PEP, TDMA with O-STBC transmission for each user at a per-user rate  $R$  would lead to an error rate at high SNR

$$P_{TDMA} \approx \left(\frac{\rho_{lt} d_{\min,T}^2}{8}\right)^{-2} = \left(\frac{4 \rho d_{\min,T}^2}{3 \cdot 8}\right)^{-2} \quad (21)$$

where  $d_{\min,T}^2$  refers to the squared minimum distance of a constellation with  $2^{2R}$  points, aiming to achieve a per-user rate  $R$  (with two symbols transmitted per O-STBC block every four channel uses). In (21),  $\rho_{lt}$  is used so as to have a fair comparison with MAT under the same consumed average power constraint. We note the difference and similarities between (21) and (18). To achieve the same per-user rate  $R$ , TDMA requires a larger constellation size than SM-encoded MAT and its performance is therefore affected by a smaller minimum distance, but SM-encoded MAT error rate slope vs. SNR is not as steep as that of TDMA due to the presence of an additional term in (18) that scales with  $\log(\rho)$ . Let us operate TDMA at the SNR  $\rho_T$  and SM-encoded MAT at the SNR  $\rho_M$ . In order to guarantee  $P_{TDMA} = P_{SM-MAT}$ ,  $\rho_T$

$$P(\mathbf{C} \rightarrow \mathbf{E}) \leq \frac{1}{b_1 b_2} \left[ \frac{1}{b_1 b_2} + \frac{(b_2 - 1)^2}{b_2^2 (b_1 - b_2)} \exp\left(\frac{1}{b_2}\right) \text{Ei}\left(\frac{-1}{b_2}\right) - \frac{(b_1 - 1)^2}{b_1^2 (b_1 - b_2)} \exp\left(\frac{1}{b_1}\right) \text{Ei}\left(\frac{-1}{b_1}\right) \right] \quad (13)$$

and  $\rho_M$  need to satisfy the relationship

$$\left( \frac{4}{3} \frac{\rho_T d_{min,T}^2}{8} \right)^{-2} = \left( \frac{\rho_M d_{min,M}^2}{8} \right)^{-2} \log \left( \frac{\rho_M d_{min,M}^2}{8} \right), \quad (22)$$

which leads to the following SNR gap  $\Delta\rho_{dB}$

$$\begin{aligned} \Delta\rho_{dB} &= 10 \log_{10}(\rho_M) - 10 \log_{10}(\rho_T) \\ &= 10 \log_{10} \left( \frac{4}{3} \frac{d_{min,T}^2}{d_{min,M}^2} \right) + \frac{10}{2} \log_{10} \log \left( \frac{\rho_M d_{min,M}^2}{8} \right) \\ &\approx 1.25 + 10 \log_{10} \left( \frac{d_{min,T}^2}{d_{min,M}^2} \right) + \frac{10}{2} \log_{10} \log \left( \frac{\rho_M d_{min,M}^2}{8} \right). \end{aligned} \quad (23)$$

The SNR gap  $\Delta\rho_{dB}$  is a simple function of the two mechanisms highlighted above: the second term in (23) is always negative and relates to the ratio of the minimum distances while the first and third terms are positive and the latter increases with the SNR. At low SNR,  $\Delta\rho_{dB}$  may be negative and indicates that SM-encoded MAT may outperform TDMA (i.e.  $\rho_T$  has to be larger than  $\rho_M$  in order to achieve the same error rate) while as the SNR increases,  $\Delta\rho_{dB}$  increases and at some point becomes positive, indicating that TDMA outperforms SM-encoded MAT. Therefore, for a fixed rate transmission, the SNR range where SM-encoded MAT is expected to outperform TDMA is concentrated at low SNR and increases as the rate increases. With a full-rate full-diversity code, MAT exhibits a larger diversity gain than TDMA and therefore always outperforms TDMA at high SNR. We recall that 1.25 dB gap originates from the fact that to operate at an SNR  $\rho$ , MAT requires to consume 4/3 more power than TDMA. Simulation results in Section VI will confirm the observations highlighted in this section. In particular, it will be shown that SM-encoded MAT is expected to exhibit some performance benefits at low SNR over TDMA only for a per-user rate above 3 bit/s/Hz.

### B. Diversity-Multiplexing Tradeoff (DMT)

In this section, we derive the optimal DMT  $(r, d^*(r))$ , as defined in [19], of the MAT scheme and the ones achievable with several space-time encoded MAT architectures with QAM constellation.  $r$  is the per-user multiplexing gain and  $d^*(r)$  is the optimal diversity gain at asymptotic high SNR. In [14], an upper bound of the asymptotic DMT (at infinite SNR) for the MAT scheme was derived. We show in Theorem 1 that this upper bound is actually the exact DMT.

**Theorem 1:** The asymptotic DMT  $(r, d^*(r))$  of the MAT scheme over i.i.d. Rayleigh fading channel is the piecewise-linear function joining the points  $(0, 3)$ ,  $(\frac{1}{3}, 1)$  and  $(\frac{2}{3}, 0)$ .

*Proof:* The proof is provided in Appendix B.  $\square$

Approximately universal codes, as defined in [22], achieve the optimal DMT at infinite SNR for any fading distribution.

The AMT strategy combined with such codes would therefore achieve the optimal SMT of Theorem 1.

**Theorem 2:** The asymptotic DMT  $(r, d(r))$  achieved by SM-encoded MAT with ML decoding and QAM constellation over i.i.d. Rayleigh fading channel is given by  $d(r) = 2 - 3r$  for  $r \in [0, \frac{2}{3}]$ .

*Proof:* The proof is provided in Appendix C.  $\square$

We observe that a simple SM is suboptimal in the range  $r \in [0, \frac{1}{3}]$ .

**Theorem 3:** The asymptotic DMT  $(r, d(r))$  achieved by O-STBC-encoded MAT with QAM constellation over i.i.d. Rayleigh fading channel is given by  $d(r) = 3(1 - 3r)$  for  $r \in [0, \frac{1}{3}]$ .

*Proof:* Given the PEP expressions in (20) at high SNR, the proof is straightforward and directly re-uses the derivations made for O-STBC in conventional Rayleigh fading MIMO channels [1], [19].  $\square$

O-STBC is clearly sub-optimal for any  $r > 0$ . At  $r = 0$  (i.e. constant rate transmission), O-STBC achieves the maximum diversity gain (i.e. 3), which was confirmed from the PEP analysis.

For comparison, a simple TDMA transmission that ignores the delayed CSIT would achieve a DMT of  $d(r) = 2(1 - 2r)$  for a per-user multiplexing gain  $r \in [0, \frac{1}{2}]$ , i.e. the line joining the points  $(0, 2)$ ,  $(\frac{1}{2}, 0)$ . Transmission using O-STBC within a coherence time for each user at a time would achieve such DMT.

It is important to recall that the DMT results are only valid at infinite SNR. While a characterization of the DMT at finite SNR is very challenging, a closer look at the proofs in the appendices highlights that the outage probability and the union bound on the error probability often scale as  $\rho^{-m} \log(\rho^{-n})$  (with  $m > 0$  and  $n > 0$ ) at high but finite SNR. While the log term does not impact the diversity gain at asymptotic high SNR, it does at finite SNR (similarly to the analysis made on the finite SNR average PEP). This explains why the PEP analysis highlights some benefits of TDMA at finite SNR even though the asymptotic DMT of TDMA is clearly lower than the optimal asymptotic DMT of MAT and that achieved with SM-encoded MAT with QAM constellation.

### C. Space-Time Code Design

From (14), focusing on the worst-case PEP, full rank codes should be designed in MAT such that

$$d_{\lambda, MAT} = \max_{\substack{\mathbf{C}, \mathbf{E} \\ \mathbf{C} \neq \mathbf{E}}} (\lambda_1 \lambda_2)^{-1} \frac{\log(\lambda_1) - \log(\lambda_2)}{\lambda_1 - \lambda_2} \quad (24)$$

is minimized, where  $\lambda_1$  and  $\lambda_2$  are the two eigenvalues of the error matrix  $\tilde{\mathbf{E}} = (\mathbf{C} - \mathbf{E})(\mathbf{C} - \mathbf{E})^H$ .

Compared to classical point-to-point i.i.d. Rayleigh fading channels, the coding gain is now a function of the quantity

$\frac{\log(\lambda_1) - \log(\lambda_2)}{\lambda_1 - \lambda_2}$ . We can write

$$\begin{aligned} & \max_{\substack{\mathbf{C}, \mathbf{E} \\ \mathbf{C} \neq \mathbf{E}}} \frac{\log(\lambda_1) - \log(\lambda_2)}{\lambda_1 - \lambda_2} \\ & \geq \max_{\substack{\mathbf{C}, \mathbf{E} \\ \mathbf{C} \neq \mathbf{E}}} \frac{2}{\lambda_1 + \lambda_2} = \frac{2}{\min_{\substack{\mathbf{C}, \mathbf{E} \\ \mathbf{C} \neq \mathbf{E}}} \|\mathbf{C} - \mathbf{E}\|_F^2}. \end{aligned} \quad (25)$$

Equality occurs whenever the minimum coding gain error matrix (i.e. the one leading to the maximum of the left-hand side of (25)) has equal eigenvalues. If equality is achieved, and assuming we aim at minimizing the worst-case PEP, codes with large minimum trace among all error matrices should be favored.

Let us further assume a linear Space-Time Block Code whose codewords write as

$$\mathbf{C} = \sum_{q=1}^Q \Phi_q \Re[c_q] + \Phi_{q+Q} \Im[c_q] \quad (26)$$

with  $\{\text{Tr}\{\Phi_q \Phi_q^H\} = T/Q\}_{q=1}^{2Q}$  [1].  $\Phi_q$  are complex basis matrices of size  $n_t \times T$ ,  $c_q$  stands for the complex information symbol,  $Q$  is the number of complex symbols  $c_q$  transmitted over a codeword,  $\Re$  and  $\Im$  stand for the real and imaginary parts. From [1] (Proposition 6.4), denoting the minimum squared Euclidean distance of the constellation used by  $d_{\min}^2$ , we know that

$$\min_{\substack{\mathbf{C}, \mathbf{E} \\ \mathbf{C} \neq \mathbf{E}}} \|\mathbf{C} - \mathbf{E}\|_F^2 \leq \frac{T}{Q} d_{\min}^2 \quad (27)$$

and equality is achieved if the basis matrices  $\{\Phi_q\}_{q=1}^{2Q}$  satisfy the conditions

$$\text{Tr}\{\Phi_q \Phi_p^H + \Phi_p \Phi_q^H\} = 0, \quad q \neq p \quad (28)$$

or equivalently

$$\mathcal{X}^T \mathcal{X} = \frac{T}{Q} \mathbf{I}_{2Q} \quad (29)$$

with

$$\mathcal{X} = \begin{bmatrix} \text{vec}\left(\begin{bmatrix} \Re[\Phi_1] \\ \Im[\Phi_1] \end{bmatrix}\right) & \cdots & \text{vec}\left(\begin{bmatrix} \Re[\Phi_{2Q}] \\ \Im[\Phi_{2Q}] \end{bmatrix}\right) \end{bmatrix}. \quad (30)$$

Combining (27) with (25), we can write

$$\max_{\substack{\mathbf{C}, \mathbf{E} \\ \mathbf{C} \neq \mathbf{E}}} \frac{\log(\lambda_1) - \log(\lambda_2)}{\lambda_1 - \lambda_2} \geq \frac{2Q}{Td_{\min}^2}. \quad (31)$$

From previous discussions, we conclude that the equality in (31) is achieved if the following two conditions are satisfied: 1)  $\mathcal{X}^T \mathcal{X} = \frac{T}{Q} \mathbf{I}_{2Q}$  and 2) the minimum coding gain error event (i.e. the one leading to the maximum of the left-hand side of (25)) is such that  $\tilde{\mathbf{E}} = \frac{Td_{\min}^2}{2Q} \mathbf{I}_2$ . We note that approximately universal codes as defined in [22] commonly satisfy condition 1 but do not satisfy condition 2 [1]. Combined with the classical min det maximization design criterion [18], those two conditions provide further insights into how to enhance space-time code designs for BC with delayed CSIT.

#### D. A Larger Number of Users and Antennas

Analysis and discussions have been limited to the two-user MAT so far. However the MAT strategy is also known for a general K-user scenario [3]. The extension of the error rate and DMT analysis to the K-user is beyond the scope of this paper. Nevertheless it is expected that the K-user MAT, similarly to the two-user scheme, will also be subject to a diversity “loss” (compared to classical Rayleigh distribution) at finite SNR owing to the double Rayleigh distribution of some of the channel coefficients.

#### IV. SPATIALLY CORRELATED FADING CHANNELS

We now extend the discussion to spatially correlated channels with any transmit correlation matrices  $\mathbf{R}_{t,1}$  and  $\mathbf{R}_{t,2}$ .

##### A. Error Rate Performance

Following the derivations in Appendix A, the average PEP of MAT is upper bounded as displayed in (32) where  $b_{k,i} = 1 + a_{k,i}$  ( $k, i = 1, 2$ ) with  $a_{k,i} = \frac{\rho}{4} \lambda_{k,i}$  and  $\lambda_{k,i}$  the  $k^{th}$  eigenvalue of the matrix  $\mathbf{R}_{t,i} \tilde{\mathbf{E}}$ .

Assuming a full rank code, i.e.  $\lambda_{k,i} > 0$  for  $k, i = 1, 2$ , the average PEP for large enough  $a_{1,2}$  and  $a_{2,2}$  can be approximated as

$$P(\mathbf{C} \rightarrow \mathbf{E}) \leq (b_{1,1} b_{2,1})^{-1} \frac{\log(b_{1,2}) - \log(b_{2,2})}{b_{1,2} - b_{2,2}} \quad (33)$$

with  $b_{1,1} b_{2,1} = \det(\mathbf{I}_2 + \rho/4 \mathbf{R}_{t,1} \tilde{\mathbf{E}})$ .

In the more restrictive condition that  $a_{k,i} \forall k, i$  are all large (i.e. high SNR),

$$P(\mathbf{C} \rightarrow \mathbf{E}) \approx \left(\frac{\rho}{4}\right)^{-3} \frac{1}{\lambda_{1,1} \lambda_{2,1}} \frac{\log(\lambda_{1,2}) - \log(\lambda_{2,2})}{\lambda_{1,2} - \lambda_{2,2}} \quad (34)$$

with  $\lambda_{1,1} \lambda_{2,1} = \det(\mathbf{R}_{t,1}) \det(\tilde{\mathbf{E}})$ . The maximum achievable diversity gain is 3. User 1's performance is a function of user 2's spatial correlation matrix  $\mathbf{R}_{t,2}$ .

With SM-encoded MAT, the rank-1 error matrix  $\tilde{\mathbf{E}}$  has a unique non-zero eigenvalue  $\lambda_{1,i}$ ,  $i = 1, 2$ . Hence  $\lambda_{2,i} = 0$ ,  $b_{2,i} = 1$ ,  $b_{1,i} = 1 + a_{1,i}$  and  $b_{1,i} - b_{2,i} = a_{1,i}$  with  $a_{1,i} = \frac{\rho}{4} \lambda_{1,i}$ . The average PEP (32) simplifies into

$$P(\mathbf{C} \rightarrow \mathbf{E}) \leq \frac{1}{b_{1,1}} \left[ \frac{1}{b_{1,2}} - \frac{a_{1,2}}{b_{1,2}^2} \exp\left(\frac{1}{b_{1,2}}\right) \text{Ei}\left(\frac{-1}{b_{1,2}}\right) \right]. \quad (35)$$

For large  $a_{1,2}$ ,  $P(\mathbf{C} \rightarrow \mathbf{E}) \leq \frac{\log(b_{1,2})}{b_{1,1} b_{1,2}}$ , leading to

$$\begin{aligned} P(\mathbf{C} \rightarrow \mathbf{E}) & \leq \log\left(1 + \frac{\rho}{4} \text{Tr}\{\mathbf{R}_{t,2} \tilde{\mathbf{E}}\}\right) \\ & \left(1 + \frac{\rho}{4} \text{Tr}\{\mathbf{R}_{t,e} \tilde{\mathbf{E}}\} + \left(\frac{\rho}{4}\right)^2 \text{Tr}\{\mathbf{R}_{t,1} \tilde{\mathbf{E}}\} \text{Tr}\{\mathbf{R}_{t,2} \tilde{\mathbf{E}}\}\right)^{-1} \end{aligned} \quad (36)$$

where  $\mathbf{R}_{t,e} = \mathbf{R}_{t,1} + \mathbf{R}_{t,2}$ . Bound (36) should be contrasted with the performance of conventional MU-MIMO and point-to-point MIMO. This will be discussed in detail in Section IV-B.

$$P(\mathbf{C} \rightarrow \mathbf{E}) \leq \frac{1}{b_{1,1}b_{2,1}} \left[ \frac{1}{b_{1,2}b_{2,2}} + \frac{(b_{2,2}-1)^2}{b_{2,2}^2(b_{1,2}-b_{2,2})} \exp\left(\frac{1}{b_{2,2}}\right) \text{Ei}\left(\frac{-1}{b_{2,2}}\right) - \frac{(b_{1,2}-1)^2}{b_{1,2}^2(b_{1,2}-b_{2,2})} \exp\left(\frac{1}{b_{1,2}}\right) \text{Ei}\left(\frac{-1}{b_{1,2}}\right) \right] \quad (32)$$

With O-STBC (Alamouti code) encoded MAT,  $\tilde{\mathbf{E}} = \alpha \mathbf{I}_2$ ,  $b_{k,i} = 1 + \frac{\rho}{4} \alpha \lambda_{k,i}$ . Assuming  $a_{1,2}$  and  $a_{2,2}$  are large enough, the average PEP in (33) simplifies as

$$P(\mathbf{C} \rightarrow \mathbf{E}) \leq \frac{\left( \det \left( \mathbf{I}_2 + \frac{\rho}{4} \alpha \mathbf{R}_{t,1} \right) \right)^{-1} \log \left( 1 + \frac{\rho}{4} \alpha \lambda_{1,2} \right) - \log \left( 1 + \frac{\rho}{4} \alpha \lambda_{2,2} \right)}{\frac{\rho}{4} \alpha (\lambda_{1,2} - \lambda_{2,2})}. \quad (37)$$

At high SNR,

$$\begin{aligned} P(\mathbf{C} \rightarrow \mathbf{E}) &\leq \left( \frac{\rho}{4} \right)^{-3} \alpha^{-3} (\det(\mathbf{R}_{t,1}))^{-1} \\ &\quad \frac{\log(\lambda_1(\mathbf{R}_{t,2})) - \log(\lambda_2(\mathbf{R}_{t,2}))}{\lambda_1(\mathbf{R}_{t,2}) - \lambda_2(\mathbf{R}_{t,2})}, \\ &= \left( \frac{\rho}{4} \right)^{-3} \alpha^{-3} (1 - |t_1|^2)^{-1} \\ &\quad \frac{\log(1 + |t_2|) - \log(1 - |t_2|)}{2|t_2|}. \end{aligned} \quad (38)$$

The higher the magnitude of the correlation coefficients  $|t_1|$  and  $|t_2|$ , the higher the error rate.

### B. MISO BC with Outdated CSIT vs. MU-MIMO vs. Point-to-Point MIMO

We can now make interesting observations about the PEP behaviour of (36) and contrasts with conventional MU-MIMO with perfect CSIT and point-to-point MIMO (single-user SM and TDMA based on O-STBC):

- *Subspace alignment:*  $\text{Tr}\{\mathbf{R}_{t,x}\tilde{\mathbf{E}}\}$  (with  $x = e, 1, 2$ ) suggests that the performance at low/medium SNR and high SNR depends on the alignment between the eigenvectors of the error matrix  $\tilde{\mathbf{E}}$  and those of the correlation matrix  $\mathbf{R}_{t,x}$  (i.e.  $\mathbf{R}_{t,e}$ ,  $\mathbf{R}_{t,1}$  and  $\mathbf{R}_{t,2}$ ), which is reminiscent of the point-to-point MIMO behaviour with Spatial Multiplexing [25]. The error probability is primarily a function of the worst-case alignment, which is known to be detrimental to Spatial Multiplexing performance [24], [25]. Similarly such subspace alignment is expected to be detrimental to the performance of SM-encoded MAT.
- *User orthogonality:*  $\text{Tr}\{\mathbf{R}_{t,e}\tilde{\mathbf{E}}\}$  is not a function of each individual transmit correlation matrix but only of their sum  $\mathbf{R}_{t,e}$ . As  $\mathbf{R}_{t,e}$  acts as an effective transmit correlation matrix,  $\mathbf{R}_{t,e}$  can appear close to an identity matrix even though each user experiences highly transmit correlated channels. Assume for instance that user 1 and user 2 experience the same magnitude of transmit correlation but different phase, i.e.  $t_1 = |t|e^{j\varphi_1}$  and  $t_2 = |t|e^{j\varphi_2}$ . For statistically orthogonal users with  $\varphi_2 - \varphi_1 = \pi$ ,  $\mathbf{R}_{t,e} = 2\mathbf{I}_2 \forall |t|$ . Hence, despite the presence of transmit correlated channels, each user experiences the same performance as if the channels were independent ( $t_1 = t_2 = 0$ ). This shows that as the phase shift  $\varphi_2 - \varphi_1$  increases and users get more and more statistically orthogonal to each other ( $\varphi_2 - \varphi_1 = \pi$ ), the performance is enhanced. This

behaviour is reminiscent of conventional MU-MIMO [1], [2], [16].

- *Transmit correlation:* Transmit correlation is known to be detrimental to point-to-point MIMO (using SM and O-STBC) with no channel state/distribution knowledge at the transmitter [24], [25]. The analysis of SM-encoded MAT highlights that the performance in spatially correlated channels approaches that of uncorrelated channels as long as  $|t_1| \approx |t_2|$  and  $\phi \approx \pi$  (i.e. user channels are statistically orthogonal), so that  $\mathbf{R}_{t,e}$  effectively behaves as if transmit correlation is zero. Otherwise, transmit correlation would be detrimental to the performance. In particular, asymmetric scenarios where one of the correlation coefficients is high and the other one low ( $|t_1| \gg |t_2|$  or inversely) would lead to a higher error rate than if both coefficients were large and  $\phi \approx \pi$ . Hence, combined with a suitable user pairing, SM-encoded MAT is expected to be less sensitive to spatial correlation than TDMA.

It is interesting to make an analogy between  $\mathbf{R}_{t,e}$  and the transmit correlation matrix of a point-to-point MIMO channel, denoted as  $\mathbf{R}_t$ . Indeed,  $\mathbf{R}_t$  can always be decomposed into the sum of two matrices  $\mathbf{R}_{t,A}$  and  $\mathbf{R}_{t,B}$  so that  $\mathbf{R}_t = \mathbf{R}_{t,A} + \mathbf{R}_{t,B}$ . This originates from decomposing all channel multipaths into two independent clusters, denoted as  $A$  and  $B$ .  $\mathbf{R}_{t,x}$ , with  $x = A, B$ , can be viewed as the transmit correlation matrix accounting for the cluster  $x$  of multipaths. Choosing statistically orthogonal users so as to decrease the off-diagonal entries of  $\mathbf{R}_{t,e}$  can therefore be viewed as an hypothetical point-to-point MIMO channel where we would have the flexibility to adjust the clusters  $A$  and  $B$  of multipaths so that the transmit correlation matrix  $\mathbf{R}_t$  is better conditioned.

Note that [12] also pointed out the presence of the quantity  $\mathbf{R}_{t,e}$  in their analysis. However it originated from the use of Jensen's inequality and was therefore only observed in an upper bound of the achievable rate. The analysis here reflects the true presence of such quantity in the PEP.

### C. User Pairing with Outdated CSIT

Observations made in previous section provide useful guidelines to appropriately pair users with outdated CSIT. Let us assume a SM-encoded MAT for simplicity. The user pairing strategy should be SNR dependent.

At *low to medium SNR*, statistically orthogonal users (characterized by  $\varphi_2 - \varphi_1 \approx \pm\pi$ ) with  $|t_1| \approx |t_2|$  should be paired together so as to experience an identity effective transmit correlation matrix  $\mathbf{R}_{t,e}$ .

At *high SNR*, on the other hand, users should be paired so as to maximize  $\text{Tr}\{\mathbf{R}_{t,1}\tilde{\mathbf{E}}\}\text{Tr}\{\mathbf{R}_{t,2}\tilde{\mathbf{E}}\}$  over all possible error matrices, i.e. users should be paired such that the dominant eigenvector of each error matrix is never aligned with the weakest eigenvector of either  $\mathbf{R}_{t,1}$  or  $\mathbf{R}_{t,2}$ . Pairing statistically orthogonal users helps decreasing the error rate as confirmed by the simulations in Section VI.



#### D. Robust Code Design

Similarly to robust code design in point-to-point MIMO channels where space-time codes and/or precoders are designed so as to avoid the detrimental alignment between eigenvectors of the error matrix and the transmit correlation matrix, precoder/code could be made robust for MISO BC with outdated CSIT. The derivations above combined with those made in [24]–[26] would provide useful guidelines to derive such precoders/codes. Nevertheless, given the similarities between MISO BC with outdated CSIT and point-to-point MIMO channels, it is expected that robust precoders designed for SM would be suitable candidates for SM-encoded MAT.

#### E. Signal Constellation Optimization

Assuming the transmit correlation matrices are known to the transmitter, we resort to an optimization of non-linear signal constellations in order to improve the performance of SM-encoded MAT. The design relies on an extension of the results obtained in [17] to the MAT transmission. In the non-linear signal constellations, the first entry of the  $2 \times 1$  codeword  $\mathbf{C}$  (assume  $T = 1$ ) is selected from a constellation  $\mathbf{S}$  (containing the symbols  $S_m$ , with  $m = 1, \dots, M_0$ ) and the second entry of  $\mathbf{C}$  is selected from a constellation  $\mathbf{Q}_m$  (containing the symbols  $Q_{mn}$ , with  $n = 1, \dots, M_1$ ), which depends on the symbol  $S_m$  chosen as the first entry of the codeword. Therefore, the second entry of  $\mathbf{C}$  is no longer independent from the first entry. The codewords write as  $\mathbf{C} = [S_m \ Q_{mn}]^T$ . Contrary to the point-to-point MIMO, the dependence between  $\mathbf{S}$  and  $\mathbf{Q}_m$  is now a function of two transmit correlation matrices rather than one. For i.i.d. channels,  $\mathbf{S}$  and  $\mathbf{Q}_m$  simply boil down to classical QAM constellations.

The average PEP of SM-encoded MAT in spatially correlated channels shows that the performance is highly determined by the quantities  $\text{Tr}\{\mathbf{R}_{t,1}\tilde{\mathbf{E}}\}$  and  $\text{Tr}\{\mathbf{R}_{t,2}\tilde{\mathbf{E}}\}$ . In particular, from (36), both users' performance in MAT are function of the quantity  $(1 + \frac{\rho}{4}\text{Tr}\{\mathbf{R}_{t,1}\tilde{\mathbf{E}}\})(1 + \frac{\rho}{4}\text{Tr}\{\mathbf{R}_{t,2}\tilde{\mathbf{E}}\})$ . In order to design the nonlinear constellations, we introduce the following objective function

$$\bar{P} \approx \frac{1}{M_0 M_1} \sum_{m=1}^{M_0} \sum_{u=1}^{M_0} \sum_{n=1}^{M_1} \sum_{v=1}^{M_1} s(S_m, S_u, Q_{mn}, Q_{uv}) \prod_{i=1}^2 \left(1 + \frac{\rho}{4} \text{Tr}\{\mathbf{R}_{t,i}\tilde{\mathbf{E}}\}\right)^{-1}, \quad (39)$$

with

$$\text{Tr}\{\mathbf{R}_{t,i}\tilde{\mathbf{E}}\} = |S_m - S_u|^2 + |Q_{mn} - Q_{uv}|^2 + 2\Re\{t_i(S_m - S_u)(Q_{mn} - Q_{uv})^*\}. \quad (40)$$

The weights, denoted as  $s$ , result from the fact that different codeword vectors  $\mathbf{C}$  and  $\mathbf{E}$  may cause a different number of symbol errors, i.e.,

$$s(c_0, e_0, c_1, e_1) = \begin{cases} 2, & c_0 - e_0 \neq 0 \text{ and } c_1 - e_1 \neq 0, \\ 1, & c_0 - e_0 = 0 \text{ or } c_1 - e_1 = 0, \\ 0, & c_0 - e_0 = 0 \text{ and } c_1 - e_1 = 0. \end{cases} \quad (41)$$

$\bar{P}$  can be seen as an estimate of the average symbol error probability (whose accuracy was demonstrated in [17] and references therein) accounting only for the term that is common to both users' performance in MAT. The problem is the selection of the signal constellations  $\mathbf{S}$  and  $\mathbf{Q}_m$  to minimize  $\bar{P}$  under an average power constraint [17]. A constrained gradient-search algorithm is used to determine the optimum constellations. Let  $\mathbf{F}$  be the vector defined as

$$\mathbf{F} = [S_1 \dots S_{M_0} \ Q_{11} \dots Q_{1M_1} \ \dots \ Q_{M_0 1} \dots Q_{M_0 M_1}]^T \quad (42)$$

and  $\mathbf{F}^k$  the vector  $\mathbf{F}$  at the  $k^{\text{th}}$  step of the algorithm. The unconstrained gradient algorithm is described by

$$\mathbf{F}^{k+1} = \mathbf{F}^k - \alpha \nabla \bar{P}(\mathbf{F}^k) \quad (43)$$

where  $\alpha$  is the step size and  $\nabla \bar{P}(\mathbf{F}^k)$  is the gradient of  $\bar{P}$  with respect to  $\mathbf{F}^k$ . The constellations are normalized at every iteration in order to account for the power constraint. The expressions of the gradients are provided in Appendix E.

#### V. OTHER STRATEGIES RELYING ON DELAYED CSIT

Aside MAT, other interesting strategies have been proposed to exploit delayed CSIT [5], [14]. Alternative MAT strategy is quite appealing as it has been shown to be a useful building block of a larger scheme [5]. Moreover it remains quite tractable and spans 3 channel uses similarly to MAT. The scheme proposed in [14] has the benefit of achieving a diversity gain of 6 (at infinite SNR), hence higher than that of MAT and Alternative MAT, but spans 24 channel uses. Thanks to the higher diversity gain, it is expected to have a better PEP performance than MAT/Alternative MAT in the infinite SNR regime but would also be subject to the effects due to the double Rayleigh distribution of some channel coefficients at finite SNR (similarly to MAT and Alternative MAT). In the sequel, we extend the past discussions to the Alternative MAT.

##### A. Space-Time Encoded Alternative MAT

Alternative MAT performs transmission in a different manner than MAT. On time slot  $t$ , Alternative MAT transmits  $\mathbf{x}_{1,t} = \sqrt{\rho}(\mathbf{c}_t + \mathbf{c}'_t)$  in coherence time 1, user 1's overheard interference  $\mathbf{x}_{2,t} = \sqrt{\rho}[\mathbf{h}_1 \mathbf{c}'_t \ 0]^T$  in coherence time 2 and user 2's overheard interference  $\mathbf{x}_{3,t} = \sqrt{\rho}[\mathbf{g}_1 \mathbf{c}_t \ 0]^T$  in coherence time 3 [5]. Hence, for a space-time encoded Alternative MAT, the equivalent system model for user 1 at time instant  $t = 1, \dots, T$  can be written as

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \end{bmatrix} = \sqrt{\rho} \begin{bmatrix} h_{1,1} & h_{1,2} \\ 0 & 0 \\ h_{3,1}g_{1,1} & h_{3,1}g_{1,2} \end{bmatrix} \mathbf{c}_t + \sqrt{\rho} \begin{bmatrix} h_{1,1} & h_{1,2} \\ h_{2,1}h_{1,1} & h_{2,1}h_{1,2} \\ 0 & 0 \end{bmatrix} \mathbf{c}'_t + \begin{bmatrix} n_{1,t} \\ n_{2,t} \\ n_{3,t} \end{bmatrix}, \quad (44)$$

and after further interference elimination

$$\tilde{\mathbf{y}}_t = \begin{bmatrix} y_{3,t} \\ h_{2,1}y_{1,t} - y_{2,t} \end{bmatrix} = \sqrt{\rho} \mathbf{H} \mathbf{c}_t + \begin{bmatrix} n_{3,t} \\ h_{2,1}n_{1,t} - n_{2,t} \end{bmatrix}, \quad (45)$$



where

$$\mathbf{H} = \begin{bmatrix} h_{3,1}g_{1,1} & h_{3,1}g_{1,2} \\ h_{2,1}h_{1,1} & h_{2,1}h_{1,2} \end{bmatrix}. \quad (46)$$

ML decoder as in (9) is performed where  $\Sigma = \text{diag}\{1, 1 + |h_{2,1}|^2\}$  is the covariance matrix of the noise vector in (45). The minimization in (9) is performed over all possible codeword vectors  $\mathbf{C}$ .

For Alternative MAT, we can define

$$\tilde{\mathbf{H}} = \Sigma^{-1/2}\mathbf{H} = \begin{bmatrix} Z & 0 \\ 0 & X \end{bmatrix} \underbrace{\begin{bmatrix} g_{1,1} & g_{1,2} \\ h_{1,1} & h_{1,2} \end{bmatrix}}_{\mathbf{H}'} \quad (47)$$

where  $Z = h_{3,1}$  and  $X = \frac{h_{2,1}}{\sqrt{1+|h_{2,1}|^2}}$ , such that the conditional PEP can be written as (12). Contrary to MAT, the average PEP of Alternative MAT is obtained by taking the expectation of the conditional PEP over the distribution of  $\mathbf{H}'$  and then over  $X$  and  $Z$ .

### B. Error Rate Performance of (encoded) Alternative MAT

Similar derivations can be made for the Alternative MAT. Let us assume first I.I.D. Rayleigh Fading Channels. Following results derived in Appendix F, the average PEP can be written as displayed in (48). At high SNR, assuming full rank code, the average PEP (48) can be approximated as

$$\begin{aligned} P(\mathbf{C} \rightarrow \mathbf{E}) &\leq \left( \frac{\log(a_1) - \log(a_2)}{a_1 - a_2} \right)^2 \\ &= \left( \frac{\rho}{4} \right)^{-2} \left( \frac{\log(\lambda_1) - \log(\lambda_2)}{\lambda_1 - \lambda_2} \right)^2. \end{aligned} \quad (49)$$

The maximum achievable diversity gain is 2 and the coding gain is proportional to  $\frac{\log(\lambda_1) - \log(\lambda_2)}{\lambda_1 - \lambda_2}$ . Comparing (14) and (49), we note the loss incurred by the Alternative MAT compared to the original MAT in terms of diversity and coding gains. The error probability of Alternative MAT is therefore expected to be significantly higher than that of MAT. Intuitively, a maximum diversity gain of 2 (rather than 3 or 4) is achieved because of the presence of  $h_{3,1}$  and  $h_{2,1}$  in both entries of respectively the first and second rows of  $\mathbf{H}$  in (46). This implies that an error is likely to occur whenever the two channel coefficients  $h_{3,1}$  and  $h_{2,1}$  are in deep fade.

From (49), we also note that if we aim at minimizing the worst-case PEP,

$$d_{\lambda, \text{AltMAT}} = \max_{\substack{\mathbf{C}, \mathbf{E} \\ \mathbf{C} \neq \mathbf{E}}} \left( \frac{\log(\lambda_1) - \log(\lambda_2)}{\lambda_1 - \lambda_2} \right)^2 \quad (50)$$

should be minimized. Hence the code design discussed in Section III-C also applies to Alternative MAT.

For SM-encoded Alternative MAT, the average PEP becomes

$$\begin{aligned} P(\mathbf{C} \rightarrow \mathbf{E}) &\leq \frac{1}{a} \exp\left(\frac{1}{a}\right) \text{Ei}\left(\frac{-1}{a}\right) \\ &\quad \left[ \frac{a}{(1+a)^2} \exp\left(\frac{1}{1+a}\right) \text{Ei}\left(\frac{-1}{1+a}\right) - \frac{1}{1+a} \right]. \end{aligned} \quad (51)$$

At high SNR,

$$P(\mathbf{C} \rightarrow \mathbf{E}) \leq \left( \frac{\log(a)}{a} \right)^2 = \left( \frac{\rho}{4} \right)^{-2} \left( \log\left(\frac{\rho}{4}\right) \right)^2. \quad (52)$$

The diversity gain at high SNR writes as  $d = 2 \left( 1 - \frac{1}{\log(a)} \right) \approx 2$ . Due to the  $(\log(\rho))^2$  term, following the discussion in Section III-A, the SNR at which Alternative MAT is outperformed by TDMA is even smaller than that of MAT.

With O-STBC-encoded Alternative MAT, the average PEP is written as

$$\begin{aligned} P(\mathbf{C} \rightarrow \mathbf{E}) &\leq \frac{1}{a^2} \left[ a + \exp\left(\frac{1}{a}\right) \text{Ei}\left(\frac{-1}{a}\right) \right] \\ &\quad \left[ \frac{b^3 - b^2 + b}{b^4} + \frac{(1-b)^2}{b^4} \exp\left(\frac{1}{b}\right) \text{Ei}\left(\frac{-1}{b}\right) \right], \end{aligned} \quad (53)$$

which leads at high SNR to

$$\begin{aligned} P(\mathbf{C} \rightarrow \mathbf{E}) &\leq \left[ \frac{1}{a} - \frac{\log(a)}{a^2} \right] \left[ \frac{1}{a} + \frac{\log(a)}{a^2} \right] \\ &= \frac{1}{a^2} - \left( \frac{\log(a)}{a^2} \right)^2 \approx \frac{1}{a^2}. \end{aligned} \quad (54)$$

Similarly to SM-encoded Alternative MAT, a diversity gain of 2 is also achieved with O-STBC.

The performance of Alternative MAT can also be evaluated in spatially correlated Rayleigh fading channels. Following Appendix F, the average PEP can be written as displayed in (55). At high SNR, assuming full rank code, the average PEP (55) can be approximated as

$$P(\mathbf{C} \rightarrow \mathbf{E}) \leq \left( \frac{\rho}{4} \right)^{-2} \prod_{i=1}^2 \frac{\log(\lambda_{1,i}) - \log(\lambda_{2,i})}{\lambda_{1,i} - \lambda_{2,i}}. \quad (56)$$

For SM, the average PEP becomes

$$\begin{aligned} P(\mathbf{C} \rightarrow \mathbf{E}) &\leq \frac{1}{a_{1,2}} \exp\left(\frac{1}{a_{1,2}}\right) \text{Ei}\left(\frac{-1}{a_{1,2}}\right) \\ &\quad \left[ \frac{a_{1,1}}{b_{1,1}^2} \exp\left(\frac{1}{b_{1,1}}\right) \text{Ei}\left(\frac{-1}{b_{1,1}}\right) - \frac{1}{b_{1,1}} \right]. \end{aligned} \quad (57)$$

At high SNR,

$$\begin{aligned} P(\mathbf{C} \rightarrow \mathbf{E}) &\leq \left( \frac{\rho}{4} \right)^{-2} \frac{\log\left(\frac{\rho}{4}\lambda_{1,1}\right) \log\left(\frac{\rho}{4}\lambda_{1,2}\right)}{\lambda_{1,1} \lambda_{1,2}}, \\ &= \left( \frac{\rho}{4} \right)^{-2} \frac{\log\left(\frac{\rho}{4}\text{Tr}\{\mathbf{R}_{t,1}\tilde{\mathbf{E}}\}\right) \log\left(\frac{\rho}{4}\text{Tr}\{\mathbf{R}_{t,2}\tilde{\mathbf{E}}\}\right)}{\text{Tr}\{\mathbf{R}_{t,1}\tilde{\mathbf{E}}\} \text{Tr}\{\mathbf{R}_{t,2}\tilde{\mathbf{E}}\}}. \end{aligned} \quad (58)$$

With O-STBC (Alamouti code), the average PEP (56) at high SNR is written as

$$\begin{aligned} P(\mathbf{C} \rightarrow \mathbf{E}) &\leq \left( \frac{\rho}{4} \right)^{-2} \alpha^{-2} \prod_{i=1}^2 \frac{\log(\lambda_1(\mathbf{R}_{t,i})/\lambda_2(\mathbf{R}_{t,i}))}{\lambda_1(\mathbf{R}_{t,i}) - \lambda_2(\mathbf{R}_{t,i})}, \\ &= \left( \frac{\rho}{4} \right)^{-2} \alpha^{-2} \prod_{i=1}^2 \frac{\log((1+|t_i|)/(1-|t_i|))}{2|t_i|}. \end{aligned} \quad (59)$$

Here also, similarly to MAT with O-STBC, the transmit correlation leads to a degradation of the PEP performance.

$$P(C \rightarrow E) \leq \left[ \frac{\exp\left(\frac{1}{a_2}\right) \text{Ei}\left(\frac{-1}{a_2}\right) - \exp\left(\frac{1}{a_1}\right) \text{Ei}\left(\frac{-1}{a_1}\right)}{a_1 - a_2} \right] \left[ \frac{1}{b_1 b_2} + \frac{(b_2 - 1)^2}{b_2^2 (b_1 - b_2)} \exp\left(\frac{1}{b_2}\right) \text{Ei}\left(\frac{-1}{b_2}\right) - \frac{(b_1 - 1)^2}{b_1^2 (b_1 - b_2)} \exp\left(\frac{1}{b_1}\right) \text{Ei}\left(\frac{-1}{b_1}\right) \right] \quad (48)$$

$$P(C \rightarrow E) \leq \left[ \frac{\exp\left(\frac{1}{a_{2,2}}\right) \text{Ei}\left(\frac{-1}{a_{2,2}}\right) - \exp\left(\frac{1}{a_{1,2}}\right) \text{Ei}\left(\frac{-1}{a_{1,2}}\right)}{a_{1,2} - a_{2,2}} \right] \left[ \frac{1}{b_{1,1} b_{2,1}} + \frac{(b_{2,1} - 1)^2}{b_{2,1}^2 (b_{1,1} - b_{2,1})} \exp\left(\frac{1}{b_{2,1}}\right) \text{Ei}\left(\frac{-1}{b_{2,1}}\right) - \frac{(b_{1,1} - 1)^2}{b_{1,1}^2 (b_{1,1} - b_{2,1})} \exp\left(\frac{1}{b_{1,1}}\right) \text{Ei}\left(\frac{-1}{b_{1,1}}\right) \right] \quad (55)$$

### C. Diversity-Multiplexing Tradeoff (DMT)

We derive the asymptotic DMT of Alternative MAT in I.I.D. Rayleigh Fading Channels.

**Theorem 4:** The asymptotic DMT  $(r, d^*(r))$  of the Alternative MAT scheme over i.i.d. Rayleigh fading channel is given by  $d^*(r) = 2 - 3r$  for  $r \in [0, \frac{2}{3}]$ , i.e. the piecewise-linear function joining the points  $(0, 2)$ ,  $(\frac{1}{3}, 1)$  and  $(\frac{2}{3}, 0)$ .

*Proof:* The proof is similar to that of Theorem 1 and is therefore omitted for brevity.  $\square$

Comparing Theorem 1 and 4, we observe that the asymptotic DMT of Alternative MAT is lower than that of MAT in the region  $r \in [0, \frac{1}{3}]$ .

**Theorem 5:** The asymptotic DMT  $(r, d(r))$  achieved by SM-encoded Alternative MAT with ML decoding and QAM constellation over i.i.d. Rayleigh fading channel is given by  $d(r) = 2 - 3r$  for  $r \in [0, \frac{2}{3}]$ .

*Proof:* The proof is provided in Appendix C.  $\square$

The DMT achieved by SM in Theorem 5 is the same as that in Theorem 2. We observe that a simple SM (i.e. no spatial encoding) is sufficient to achieve the optimal DMT with Alternative MAT but is suboptimal in the range  $r \in [0, \frac{1}{3}]$  with MAT.

**Theorem 6:** The asymptotic DMT  $(r, d(r))$  achieved by O-STBC-encoded Alternative MAT with QAM constellation over i.i.d. Rayleigh fading channel is given by  $d(r) = 2(1 - 3r)$  for  $r \in [0, \frac{1}{3}]$ .

*Proof:* Given the PEP expressions in (54) at high SNR, the proof is straightforward and directly re-uses the derivations made for O-STBC in conventional Rayleigh fading MIMO channels [1], [19].  $\square$

The DMT achieved in Theorem 6 is lower than that obtained in Theorem 3. Figure 1 summarizes all DMT at infinite SNR derived in Theorem 1 to 6.

## VI. PERFORMANCE EVALUATIONS

Figure 2 illustrates the BER performance of MAT and Alternative MAT with SM and a full-rate full-diversity approximately universal space-time code, denoted as “Dayal” [28] over i.i.d. Rayleigh fading channels. QPSK is assumed so that it corresponds to a 4/3-bit/s/Hz transmission per user with “MAT-SM”, “Alt MAT-SM”, “MAT-Dayal”, “Alt MAT-Dayal”. As a baseline, we also display the performance of TDMA based on O-STBC and 8PSK over a conventional point-to-point MISO i.i.d. Rayleigh slow fading channels with two

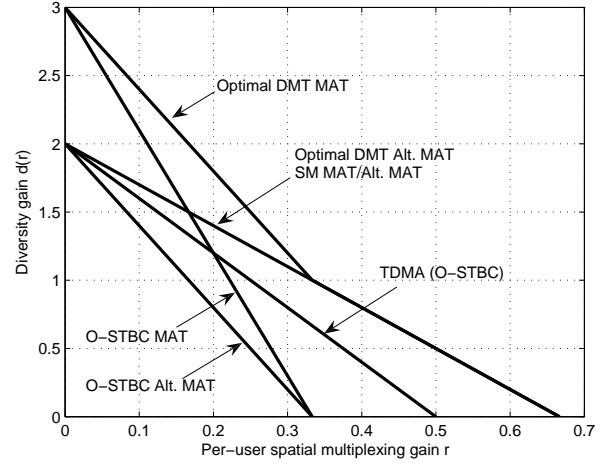


Fig. 1. Optimal DMT of MAT and Alternative MAT and achievable DMT with SM and O-STBC encoded MAT/Alternative MAT and TDMA.

transmit antennas, leading to a 3/2-bit/s/Hz transmission per user. Figure 3 extends the comparison between SM-encoded MAT and Alternative MAT and TDMA at higher rates. Note that the displayed SNR on the x-axis is  $\rho$ . Hence for a SNR of  $\rho$ , TDMA is allocated a power  $4/3\rho$  in order to keep the total average transmit power the same for MAT and TDMA. The behaviour follows the observations made from the analytical results on diversity and coding gains of TDMA with O-STBC and MAT/Alt MAT with SM/full rank codes (Section III-A). It also confirms that TDMA has a larger diversity gain than SM-encoded MAT and Alternative MAT but is impacted by the use of larger constellation sizes. Namely, we make the following observations: 1) At low rate (3 bit/s/Hz per user and below), MAT outperforms TDMA at high SNR if it is concatenated with a full-rate full-diversity space-time code (see Figure 2), 2) SM-encoded MAT does not show any significant benefit over TDMA in the simulated rate and SNR range (Figures 2 and 3), 3) Alternative MAT does not show any benefit over TDMA in the simulated rate and SNR range (Figures 2 and 3). At even higher rate (above 3 bit/s/Hz per user), SM-encoded MAT (i.e. without requiring any additional encoding in the spatial domain) is expected to start showing some meaningful performance benefits over TDMA at low SNR.

In Figure 4, the BER performance of user 1 with QPSK-based SM-encoded MAT (i.e.  $R=4/3$  bit/s/Hz per user) in

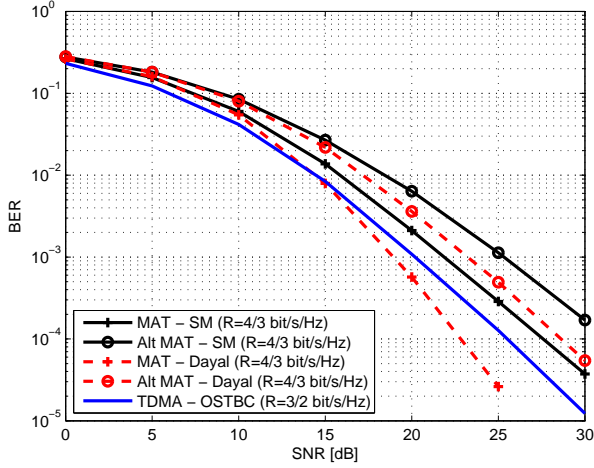


Fig. 2. Average Bit Error Rate (BER) of SM and “Dayal” encoded MAT and Alternative MAT with a per-user rate  $R=4/3$  bit/s/Hz (corresponding to using QPSK) and comparison with TDMA with a per-user rate  $R=3/2$  bit/s/Hz (O-STBC with 8PSK).

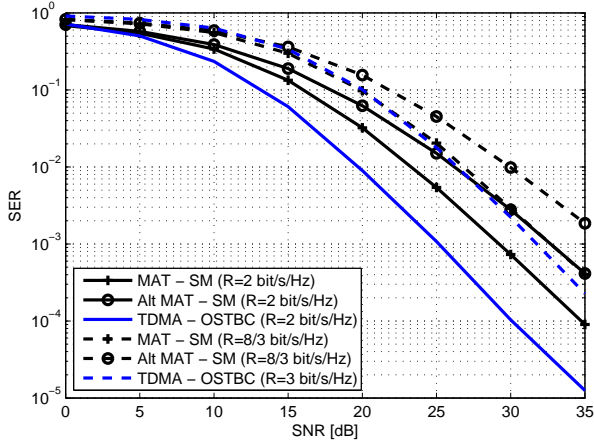


Fig. 3. Average Symbol Error Rate (SER) of SM-encoded MAT and Alternative MAT with a per-user rate  $R=2$  bit/s/Hz (corresponding to using 8PSK) and  $R=8/3$  bit/s/Hz (corresponding to using 16-QAM) and comparison with TDMA with a per-user rate  $R=2$  bit/s/Hz (O-STBC with 16-QAM) and  $R=3$  bit/s/Hz (O-STBC with 64-QAM).

spatially correlated channels is displayed for various pairs of the transmit correlation coefficients  $(t_1, t_2)$ . The performance of TDMA with OSTBC at a rate per user of  $R=3/2$  bit/s/Hz is also provided. In the presence of high transmit correlation, the phase of the correlation coefficient is indicative of the location of the user w.r.t. the transmit array as  $\varphi \approx 2\pi d/\lambda \cos \theta$  with  $d$  the inter-element spacing and  $\theta$  the angle of departure (taken w.r.t. the array axis). Increasing the transmit correlation coefficients phase shift (i.e.  $\phi = \varphi_2 - \varphi_1$ ) makes the users' channels more statistically orthogonal. In the evaluations,  $\psi$  is random and uniformly distributed within  $[0, 2\pi]$  such that the scenario  $(t_1, t_2) = (0.99e^{j\psi}, 0.99e^{j(\psi+\phi)})$  refers to the case where  $|t_1| = |t_2| = 0.99$ ,  $\varphi_1 = \psi$  is randomly distributed and  $\varphi_2 = \varphi_1 + \phi$ . We make the following observations, inline with the analytical derivations:

- 1) The performance improves as users get statistically

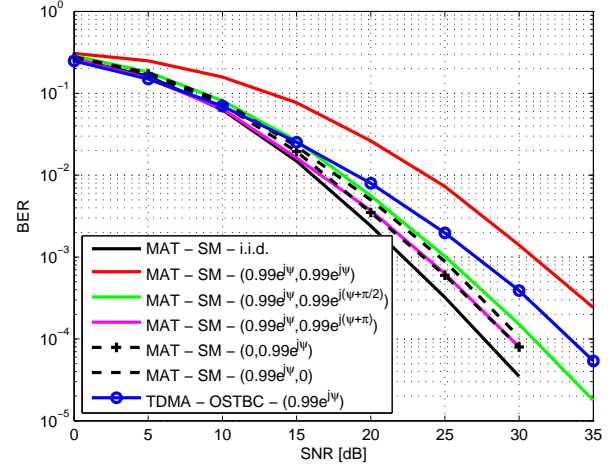


Fig. 4. Average Bit Error Rate (BER) of SM-encoded MAT (with  $R=4/3$  bit/s/Hz per user) over spatially correlated channels  $(t_1, t_2)$ . Phase  $\psi$  is random and uniformly distributed within  $[0, 2\pi]$ . Comparison with TDMA - OSTBC ( $R=3/2$  bit/s/Hz) is also provided.

orthogonal to each other. By decreasing order of BER, we have  $\phi$  equal to  $0, \pi/2, \pi$ .

- 2) I.i.d. channels and spatially correlated channels with statistically orthogonal users  $(0.99e^{j\psi}, 0.99e^{j(\psi+\pi)})$  lead to similar performance at low to medium SNR (curves are superposed up to 15dB).
- 3) Pairing two statistically orthogonal users with similar magnitudes of the transmit correlation coefficients outperforms pairing two users with asymmetric spatial correlation (one with large spatial correlation and the other one with low spatial correlation). This can be seen by comparing  $(0.99e^{j\psi}, 0.99e^{j(\psi+\pi)})$  with  $(0, 0.99e^{j\psi})$  and  $(0.99e^{j\psi}, 0)$ . Hence, interestingly, high transmit spatial correlation can lead to a better performance than low transmit spatial correlation. This contrasts with the conventional impact of transmit correlation on error rate performance in point-to-point channels [25].
- 4) MAT is shown to be less sensitive to spatial correlation than TDMA - OSTBC. Indeed, SM-encoded MAT outperforms TDMA in spatially correlated channels with proper user pairing (e.g.  $\phi = \pi/2, \pi$ ). This contrasts with the behavior on i.i.d. channels where SM-encoded MAT is outperformed by TDMA at the same rates (see Figure 2).

We now investigate the performance gain of non-linear signal constellations  $\mathbf{S}$  and  $\{\mathbf{Q}_m\}$  designed to replace conventional QPSK in a 4/3-bit/s/Hz transmission. They have been optimized for several pairs of  $(t_1, t_2)$ , whereas  $\rho/4$  is chosen equal to 20dB. A large number of initial conditions have been tested, and less than 1000 iterations were necessary to converge to an optimum. Several values of  $\alpha$  in (43) have been considered, depending on the speed of convergence of the algorithm. The constellations that have been used in Figure 5 are those that provide the minimum  $\bar{P}$  among the encountered local optimal constellations. In the configuration  $(0.95, 0.95e^{j\pi})$ , optimized constellations get closer to QPSK

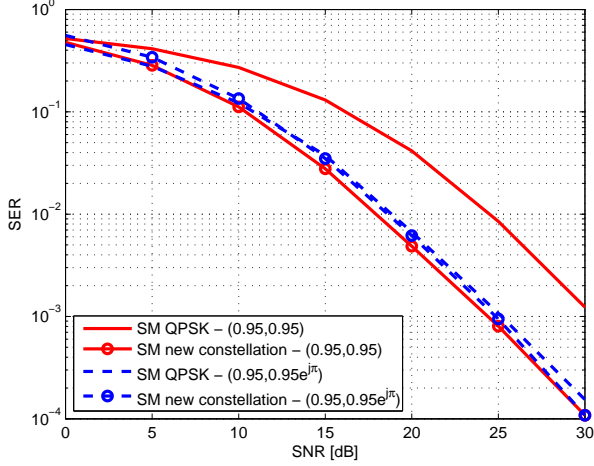


Fig. 5. Average Symbol Error Rate (SER) of SM-encoded MAT over spatially correlated channels  $(t_1, t_2)$  with QPSK and optimized constellations.

constellations as  $\mathbf{R}_{t,e}$  gets closer to a scaled identity matrix. Contrary to the QPSK constellations in Figure 4, the adaptive non-linear constellations make the performance almost insensitive to the phase shift  $\varphi_2 - \varphi_1$ . Indeed, curves are almost superposed to each other at high SNR (recall the optimization is made for  $\rho/4$  equal to 20dB). Hence the transmitter does not have to worry about user orthogonality and could schedule any pair of users irrespectively of their orthogonality.

## VII. CONCLUSIONS

The error rate performance of space-time encoded MISO BC with delayed CSIT has been analyzed. The paper provides new insights into the diversity-multiplexing tradeoff achievable with various transmissions strategies and derives novel space-time code designs, signal constellations and user pairing strategies for multi-user communications in the presence of outdated CSIT. The effect of transmit spatial correlation on the performance is also highlighted and the performance behaviour is contrasted with conventional MU-MISO with perfect CSIT and point-to-point MIMO channels. The paper contrasts with the common DoF analysis and provides new insights into the actual error rate performance benefits over TDMA. The main takeaway message is that in i.i.d. Rayleigh fading channels, MAT is not always a better alternative than TDMA. Benefits are obtained over TDMA (based on O-STBC) only at very high rate or once concatenated with a full-rate full-diversity space-time code. Alternative MAT is even less competitive than MAT and is outperformed by TDMA over a wide SNR and rate range. In spatially correlated channels, MAT with a suitable user pairing strategy is shown to be less sensitive to transmit spatial correlation than TDMA.

### APPENDIX A

#### AVERAGE PEP OF SPACE-TIME ENCODED MAT

The average PEP is obtained by taking the expectation of the conditional PEP (12) over the channel distribution. The expectation can be computed in two steps: first by taking

the expectation over the distribution of  $\mathbf{H}'$  and then over the distribution of  $X$ . The first step writes as

$$P(\mathbf{C} \rightarrow \mathbf{E} | X) = \mathcal{E}_{\mathbf{H}'} \left\{ P(\mathbf{C} \rightarrow \mathbf{E} | \tilde{\mathbf{H}}) \right\} = \left( \det \left( \mathbf{I}_4 + \frac{\rho}{4} \mathbf{R} [\mathbf{I}_2 \otimes \tilde{\mathbf{E}}] \right) \right)^{-1} \quad (60)$$

where

$$\mathbf{R} = \mathcal{E} \left\{ \text{vec}(\tilde{\mathbf{H}}^H) \text{vec}(\tilde{\mathbf{H}}^H)^H \right\} = \text{diag} \left\{ \mathbf{R}_{t,1}, |X|^2 \mathbf{R}_{t,2} \right\} \quad (61)$$

and  $\tilde{\mathbf{E}} = (\mathbf{C} - \mathbf{E})(\mathbf{C} - \mathbf{E})^H$  is the error matrix. The component block matrices  $\mathbf{R}_{t,1}$  and  $\mathbf{R}_{t,2}$  are defined as the transmit correlation matrices for receiver 1 and 2 as in (3).

Expression (60) is obtained from the classical derivation of average PEP over MIMO spatially correlated Rayleigh fading channels [1], [25]. We can further expand as

$$P(\mathbf{C} \rightarrow \mathbf{E} | X) = \left( \det \left( \mathbf{I}_2 + \frac{\rho}{4} \mathbf{R}_{t,1} \tilde{\mathbf{E}} \right) \right)^{-1} \left( \det \left( \mathbf{I}_2 + \frac{\rho}{4} |X|^2 \mathbf{R}_{t,2} \tilde{\mathbf{E}} \right) \right)^{-1} = b_{1,1}^{-1} b_{2,1}^{-1} \left( 1 + a_{1,2} |X|^2 \right)^{-1} \left( 1 + a_{2,2} |X|^2 \right)^{-1} \quad (62)$$

where  $b_{k,i} = 1 + a_{k,i}$  ( $k, i = 1, 2$ ) with  $a_{k,i} = \frac{\rho}{4} \lambda_{k,i}$  and  $\lambda_{k,i}$  the  $k^{th}$  eigenvalue of the matrix  $\mathbf{R}_{t,i} \tilde{\mathbf{E}}$ . We can moreover write  $|X|^2 = \frac{Y}{1+Y}$  where  $Y = |h_{3,1}|^2$  is exponentially distributed. In the particular case of i.i.d. Rayleigh fading,  $b_{k,1} = b_{k,2} = b_k = 1 + a_k$  ( $k = 1, 2$ ) with  $a_k = \frac{\rho}{4} \lambda_k$  and  $\lambda_k$  the  $k^{th}$  eigenvalue of the error matrix  $\tilde{\mathbf{E}}$ .

As a second step, we take the expectation over the distribution of  $X$  (or equivalently  $Y$ ). We write

$$P(\mathbf{C} \rightarrow \mathbf{E}) = \mathcal{E}_X \{ P(\mathbf{C} \rightarrow \mathbf{E} | X) \} = b_{1,1}^{-1} b_{2,1}^{-1} \int_0^\infty \left( 1 + a_{1,2} \frac{y}{1+y} \right)^{-1} \left( 1 + a_{2,2} \frac{y}{1+y} \right)^{-1} e^{-y} dy. \quad (63)$$

Making use of partial fraction expansion and Table of Integrals [27], the integral in (63) can be solved and we obtain (13) in i.i.d. Rayleigh fading channels and (32) in spatially correlated Rayleigh fading channels.  $\text{Ei}(x) = \int_{-\infty}^x \frac{e^t}{t} dt$  in (13) and (32) is the exponential integral whose Taylor series for real argument write as  $\text{Ei}(x) = \gamma + \log|x| + \sum_{i=1}^\infty \frac{x^i}{i!}$  where  $\gamma$  is the Euler-Mascheroni constant.

At high SNR, assuming a full-rank code,  $a_{k,i} \gg 0$  and  $b_{k,i} \gg 0 \forall k, i$  and for small argument  $x$ ,  $\text{Ei}(x) \approx \gamma + \log|x|$ . Applying those simplifications at high SNR to (13) and (32) and keeping only the dominant terms lead to the approximations (14) and (34). The PEP expressions for SM and O-STBC in i.i.d. channels (15) and (19) and in spatially correlated channels (35) are obtained from (13) and (32) by directly plugging the corresponding parameters. In i.i.d. channels, the PEP for O-STBC (for which  $b_1 = b_2 = b$ ) involves the computation of  $\lim_{b_1 \rightarrow b_2} \frac{f(b_1) - f(b_2)}{b_1 - b_2}$  with  $f(x) = \frac{(x-1)^2}{x^2} e^{1/x} \text{Ei}(\frac{-1}{x})$ . This can be done by noting that  $\frac{\partial \text{Ei}(x)}{\partial x} = \frac{e^x}{x}$ .



## APPENDIX B PROOF OF THEOREM 1

Following [14] and the system model (8), we can write the outage probability

$$P_{\text{out}}(r) \doteq P\left(\log_2 \det\left(\mathbf{I}_2 + \frac{\rho}{2} \mathbf{H} \mathbf{H}^H\right) < 3R\right) \quad (64)$$

where  $R = r \log_2 \rho$ . Making use of the QR decomposition  $\mathbf{H} = \mathbf{Q} \mathbf{R}$  and  $\mathbf{H}' = \mathbf{Q}' \mathbf{R}'$  with  $(i, j)$  entry of  $\mathbf{R}$  (resp.  $\mathbf{R}'$ ) denoted as  $r_{ij}$  ( $i, j = 1, 2$ ) and  $r_{21} = 0$  (resp.  $r'_{ij}$  with  $r'_{21} = 0$ ), we get

$$P_{\text{out}}(r) \doteq P\left(1 + \frac{\rho}{2} \|\mathbf{H}\|_F^2 + \left(\frac{\rho}{2}\right)^2 |\det(\mathbf{H})|^2 < \rho^{3r}\right) \quad (65)$$

where  $\|\mathbf{H}\|_F^2 = r_{11}^2 + |r_{12}|^2 + r_{22}^2$  and  $|\det(\mathbf{H})|^2 = r_{11}^2 r_{22}^2$ . Since  $\det(\mathbf{H}) = h_{3,1} \det(\mathbf{H}')$ , we also have  $|\det(\mathbf{H})|^2 = |h_{3,1}|^2 r_{11}'^2 r_{22}'^2$ .

For  $\frac{1}{3} \leq r \leq \frac{2}{3}$ ,  $1 + \frac{\rho}{2} \|\mathbf{H}\|_F^2 \ll \rho^{3r}$  such that

$$P_{\text{out}}(r) \doteq P\left(\left(\frac{\rho}{2}\right)^2 |h_{3,1}|^2 r_{11}'^2 r_{22}'^2 < \rho^{3r}\right). \quad (66)$$

Similarly to [20], we can now use the fact that the dominant error event is given by  $r_{11}'^2 < 1$  and  $|h_{3,1}|^2 r_{22}'^2 < \rho^{3r-2}$  because  $r_{11}'^2$  is likely to be larger than  $|h_{3,1}|^2 r_{22}'^2$ . Indeed  $r_{11}'^2 \sim \chi_4^2$  while  $|h_{3,1}|^2 r_{22}'^2$  writes as the product of two independent  $\chi_2^2$ .

Assuming  $X, Y \sim \chi_2^2$  and independent, we can compute the CDF of  $U = XY$  as

$$P[U \leq \epsilon] = 1 - \int_0^\infty e^{-\left(\frac{\epsilon}{v} + v\right)} dv = 1 - 2\sqrt{\epsilon} K_1(2\sqrt{\epsilon}) \quad (67)$$

where  $K_1(x)$  is the modified Bessel function of the second kind. The second equality comes from [27]. Assuming small  $x$ , from the series expansions in [21],  $x K_1(x) \approx 1 + x^2 \log(x/2)$ . Hence, for small  $\epsilon$ ,  $P[U \leq \epsilon] \approx -2\epsilon \log(\epsilon)$ .

The outage probability then writes at large SNR as

$$\begin{aligned} P_{\text{out}}(r) &\doteq P\left(r_{11}'^2 < 1 \text{ and } |h_{3,1}|^2 r_{22}'^2 < \rho^{3r-2}\right) \\ &\doteq -2\rho^{3r-2} \log(\rho^{3r-2}) \doteq \rho^{-(2-3r)}. \end{aligned} \quad (68)$$

Therefore,  $d^*(r) = 2 - 3r$  for  $\frac{1}{3} \leq r \leq \frac{2}{3}$ . Hence the piecewise-linear function joining  $(\frac{1}{3}, 1)$  and  $(\frac{2}{3}, 0)$ .

For  $0 \leq r \leq \frac{1}{3}$ ,

$$\begin{aligned} P_{\text{out}}(r) &\doteq P\left(\frac{\rho}{2} \left(|h_{3,1}|^2 [ |g_{1,1}|^2 + |g_{1,2}|^2 ] + |h_{1,1}|^2 + |h_{1,2}|^2 \right) \left( \left(\frac{\rho}{2}\right)^2 |h_{3,1}|^2 r_{11}'^2 r_{22}'^2 < \rho^{3r} \right)\right) \quad (69) \end{aligned}$$

where  $A = \log\left(\frac{\rho^{1-3r/2}}{4}\right)$ ,  $B = \log C$  and  $C = |i_1|^2 + \dots + |i_4|^2$ . Following [22],

and we can conclude from the first order term that the dominant outage events are such that  $|h_{1,1}|^2 < \rho^{3r-1}$ ,  $|h_{1,2}|^2 < \rho^{3r-1}$ ,  $|h_{3,1}|^2 < \rho^{3r-1}$ ,  $|g_{1,1}|^2 < 1$  and  $|g_{1,2}|^2 < 1$ . The last three inequalities come from the fact that  $|h_{3,1}|^2 [ |g_{1,1}|^2 + |g_{1,2}|^2 ]$  is likely to be smaller than  $\rho^{3r-1}$  whenever  $|h_{3,1}|^2 < \rho^{3r-1}$  because  $|h_{3,1}|^2 \sim \chi_2^2$  while  $|g_{1,1}|^2 + |g_{1,2}|^2 \sim \chi_4^2$ . Those inequalities imply that  $|\det(\mathbf{H}')|^2 = r_{11}'^2 r_{22}'^2 < \rho^{3r-1}$ . Therefore, in order to guarantee  $|h_{3,1}|^2 r_{11}'^2 r_{22}'^2 < \rho^{3r-2}$  in the

second order term of (69),  $|h_{3,1}|^2$  should be even smaller, less than  $\rho^{-1}$ . The outage probability writes as

$$\begin{aligned} P_{\text{out}}(r) &\doteq P\left(|h_{3,1}|^2 < \rho^{-1} \text{ and } |h_{1,1}|^2 < \rho^{3r-1} \text{ and } |h_{1,2}|^2 < \rho^{3r-1} \text{ and } |g_{1,1}|^2 < 1 \text{ and } |g_{1,2}|^2 < 1\right) \\ &\doteq \rho^{-1} \rho^{3r-1} \rho^{3r-1} \doteq \rho^{-(3-6r)}. \end{aligned} \quad (70)$$

Hence,  $d^*(r) = 3 - 6r$  for  $0 \leq r \leq \frac{1}{3}$ , i.e. the piecewise-linear function joining  $(0, 3)$  and  $(\frac{1}{3}, 1)$ . From (70), the dominant outage event occurs whenever the three channel coefficients  $h_{3,1}$ ,  $h_{1,1}$  and  $h_{1,2}$  are in deep fade.  $g_{1,1}$  and  $g_{1,2}$  do not influence the DMT as both are multiplied by  $h_{3,1}$  in (6).

## APPENDIX C PROOF OF THEOREM 2 AND 5

From (16) and (52), the average PEP with SM at high SNR writes as  $P(\mathbf{C} \rightarrow \mathbf{E}) \leq \frac{(\log(a))^m}{a^2}$  with  $m = 1$  for MAT and  $m = 2$  for Alternative MAT. Assuming without loss of generality  $\mathbf{E} = \mathbf{0}$ , we can bound the overall error probability using the union bound

$$\begin{aligned} P_e &\leq \sum_{\mathbf{C} \neq \mathbf{0}} P(\mathbf{C} \rightarrow \mathbf{0}) \\ &\leq \sum_{\mathbf{C} \neq \mathbf{0}} \left(\frac{\rho}{4}\right)^{-2} \left(\sum_{q=1}^2 |c_q|^2\right)^{-2} \left(\log\left(\frac{\rho}{4} \sum_{q=1}^2 |c_q|^2\right)\right)^m. \end{aligned} \quad (71)$$

Assuming that the same data rate  $R_q = 3R/2$  is assigned to both streams (with  $R = r \log \rho$ ), symbols  $c_q$  are chosen from a QAM constellation carrying  $2^{R_q/2}$  symbols per dimension. For a unit average energy, the minimum distance  $d_{\min}$  between two points of such a constellation is of the order  $1/2^{R_q/2}$ . We write  $c_q = (i_{2q-1} + j i_{2q}) d_{\min}$  where  $i_x \in \mathbb{Z}$  ( $x = 1, \dots, 4$ ). The minimum squared distance of the constellation is of the order  $d_{\min}^2 = 2^{-3R/2} = \rho^{-3r/2}$  and we can further bound  $P_e$  as

$$\begin{aligned} P_e &\leq \sum_{\mathbf{C} \neq \mathbf{0}} 4^2 \rho^{-(2-3r)} \frac{\left(\log\left(\frac{\rho^{1-3r/2}}{4} \sum_{x=1}^4 |i_x|^2\right)\right)^m}{\left(\sum_{x=1}^4 |i_x|^2\right)^2} \\ &= 4^2 \rho^{-(2-3r)} \sum_{(i_1, \dots, i_4) \neq \mathbf{0}} \frac{A^m + B^m + 2(m-1)AB}{C^2} \end{aligned} \quad (72)$$

$$\sum_{(i_1, \dots, i_4) \neq \mathbf{0}} \frac{1}{C^2} \doteq 1. \quad (73)$$

Similarly, from Appendix D, we can compute

$$\sum_{(i_1, \dots, i_4) \neq \mathbf{0}} \frac{B^m}{C^2} \doteq 1. \quad (74)$$

This leads to

$$\begin{aligned} P_e &\leq \rho^{-(2-3r)} \left( \left(\log\left(\rho^{1-3r/2}\right)\right)^m + (m-1) \log\left(\rho^{1-3r/2}\right) \right) \\ &\doteq \rho^{-(2-3r)}. \end{aligned} \quad (75)$$

Hence, the DMT achieved by SM-encoded MAT and Alternative MAT is given by  $d(r) = 2 - 3r$  for  $r \in [0, \frac{2}{3}]$ .

#### APPENDIX D PROOF OF (74)

Similarly to [22], the summation is split over vectors with nonzero components and is successively upper bounded using several inequalities. In the sequel,  $S$  denotes a subset of  $\{1, 2, 3, 4\}$ . We write the following inequalities

$$\begin{aligned}
& \sum_{(i_1, \dots, i_4) \neq \mathbf{0}} \frac{B^m}{C^2} \\
&= \sum_S \sum_{i_x \neq 0: x \in S} \frac{\left(\log(|i_1|^2 + \dots + |i_4|^2)\right)^m}{\left(|i_1|^2 + \dots + |i_4|^2\right)^2} \\
&= \sum_S \sum_{i_x \neq 0: x \in S} \frac{\left(\log\left(\sum_{x \in S} |i_x|^2\right)\right)^m}{\left(\sum_{x \in S} |i_x|^2\right)^2} \\
&\stackrel{(a)}{\leq} \sum_S \sum_{i_x \neq 0: x \in S} \frac{\left(\sum_{x \in S} |i_x|^2 \log(|i_x|^2)\right)^m}{\left(\sum_{x \in S} |i_x|^2\right)^{2+m}} \\
&\stackrel{(b)}{\leq} \sum_S \sum_{i_x \neq 0: x \in S} \frac{\left(\sum_{x \in S} |i_x|^2 \log(|i_x|^2)\right)^m}{\prod_{x \in S} |i_x|^{2(2+m)/|S|}} \\
&\leq \sum_S \sum_{i_x \neq 0: x \in S} \left[ \sum_{x \in S} \frac{\left(\log(|i_x|^2)\right)^m}{|i_x|^{4/|S|}} \right. \\
&\quad \left. + (m-1) \sum_{\substack{x_1, x_2 \in S \\ x_1 \neq x_2}} \frac{\log(|i_{x_1}|^2) \log(|i_{x_2}|^2)}{|i_{x_1}|^{2(1+m)/|S|} |i_{x_2}|^{2(1+m)/|S|}} \right] \\
&\leq \sum_S \sum_{i_x \neq 0: x \in S} \left[ \sum_{x \in S} \frac{\left(\log(|i_x|^2)\right)^m}{|i_x|^{4/|S|}} \right. \\
&\quad \left. + (m-1) \sum_{\substack{x_1, x_2 \in S \\ x_1 \neq x_2}} \frac{1}{|i_{x_1}|^{2m/|S|} |i_{x_2}|^{2m/|S|}} \right] \\
&\leq \sum_S \sum_{i_x \neq 0: x \in S} \left[ \sum_{x \in S} \frac{\left(\log(|i_x|^2)\right)^m}{|i_x|^{4/|S|}} \right. \\
&\quad \left. + (m-1)(|S|-1) \sum_{x \in S} \frac{1}{|i_x|^{2m/|S|}} \right] \\
&= \sum_S |S| \left[ \sum_{i_1 \neq 0} \frac{\left(\log(|i_1|^2)\right)^m}{|i_1|^{4/|S|}} \right. \\
&\quad \left. + (m-1)(|S|-1) \sum_{i_1 \neq 0} \frac{1}{|i_1|^{2m/|S|}} \right] \quad (76)
\end{aligned}$$

where (a) results from the application of the log sum inequality [23] and (b) from the inequality of arithmetic and geometric means. Since  $|S| \leq 4$ , and recalling that  $|i_1|$  grows with the

SNR,

$$\sum_{i_1 \neq 0} \frac{\left(\log(|i_1|^2)\right)^m}{|i_1|^{4/|S|}} \doteq 1 \quad \text{and} \quad \sum_{i_1 \neq 0} \frac{1}{|i_1|^{4/|S|}} \doteq 1,$$

and we finally obtain  $\sum_{(i_1, \dots, i_4) \neq \mathbf{0}} \frac{B^m}{C^2} \doteq 1$ .

#### APPENDIX E ANALYTICAL EXPRESSION OF THE GRADIENTS

The gradients of  $\bar{P}$  with respect to  $S_k$  and  $Q_{xk}$  are respectively expressed as displayed in (77) and (78), where

$$\begin{aligned}
A &= |S_k - S_u|^2 + |Q_{kn} - Q_{uv}|^2 \\
&\quad + 2\Re\{t_1(S_k - S_u)(Q_{kn} - Q_{uv})^*\}, \\
B &= |S_k - S_u|^2 + |Q_{kn} - Q_{uv}|^2 \\
&\quad + 2\Re\{t_2(S_k - S_u)(Q_{kn} - Q_{uv})^*\}, \\
C &= |S_x - S_u|^2 + |Q_{xk} - Q_{uv}|^2 \\
&\quad + 2\Re\{t_1(S_x - S_u)(Q_{xk} - Q_{uv})^*\}, \\
D &= |S_x - S_u|^2 + |Q_{xk} - Q_{uv}|^2 \\
&\quad + 2\Re\{t_2(S_x - S_u)(Q_{xk} - Q_{uv})^*\}. \quad (79)
\end{aligned}$$

#### APPENDIX F AVERAGE PEP OF SPACE-TIME ENCODED ALTERNATIVE MAT

Taking the expectation of the conditional pairwise error probability (PEP)  $P(\mathbf{C} \rightarrow \mathbf{E} | \tilde{\mathbf{H}})$  (12) over the distribution of  $\mathbf{H}'$  assuming  $Z$  and  $X$  fixed leads to

$$\begin{aligned}
P(\mathbf{C} \rightarrow \mathbf{E} | X, Z) &= \mathcal{E}_{\mathbf{H}'} \left\{ P(\mathbf{C} \rightarrow \mathbf{E} | \tilde{\mathbf{H}}) \right\} \\
&= \left( \det \left( \mathbf{I}_4 + \frac{\rho}{4} \mathbf{R} [\mathbf{I}_2 \otimes \tilde{\mathbf{E}}] \right) \right)^{-1} \quad (80)
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{R} &= \mathcal{E} \left\{ \text{vec}(\tilde{\mathbf{H}}^H) \text{vec}(\tilde{\mathbf{H}}^H)^H \right\} \\
&= \text{diag} \left\{ |Z|^2 \mathbf{R}_{t,2}, |X|^2 \mathbf{R}_{t,1} \right\} \quad (81)
\end{aligned}$$

with the transmit correlation matrices defined as in (3).

Alternatively, we can write the conditional PEP on  $X$  and  $Z$  as in (82) where  $a_{k,i}$  and  $b_{k,i}$  are defined as in Appendix A. Integrating over the distribution of  $X$  and  $Z$ , we obtain

$$\begin{aligned}
P(\mathbf{C} \rightarrow \mathbf{E}) &= \int_0^\infty \prod_{k=1}^2 (1 + a_{k,2}y)^{-1} e^{-y} dy \\
&\quad \int_0^\infty \prod_{k=1}^2 \left( 1 + a_{k,1} \frac{y}{1+y} \right)^{-1} e^{-y} dy. \quad (83)
\end{aligned}$$

Making use of partial fraction expansion and Table of Integrals [27], integrals can be solved and we obtain (48) in i.i.d. channels and (55) in spatially correlated channels.

$$g_{S_k} = \nabla_{S_k} \bar{P} = \frac{-2}{M_0 M_1} \sum_{u \neq k} \sum_{n=1}^{M_1} \sum_{v=1}^{M_1} \frac{\frac{\rho}{4} (2(S_k - S_u) + 2t_1^* (Q_{kn} - Q_{uv})) s(S_k, Q_{kn}, S_u, Q_{uv})}{(1 + \frac{\rho}{4} A)^2 (1 + \frac{\rho}{4} B)} + \frac{\frac{\rho}{4} (2(S_k - S_u) + 2t_2^* (Q_{kn} - Q_{uv})) s(S_k, Q_{kn}, S_u, Q_{uv})}{(1 + \frac{\rho}{4} A) (1 + \frac{\rho}{4} B)^2} \quad (77)$$

$$g_{Q_{xk}} = \nabla_{Q_{xk}} \bar{P} = \frac{-2}{M_0 M_1} \sum_{u=1}^{M_0} \sum_{v=1}^{M_1} \frac{\frac{\rho}{4} (2(Q_{xk} - Q_{uv}) + 2t_1 (S_x - S_u)) s(S_x, Q_{xk}, S_u, Q_{uv})}{(1 + \frac{\rho}{4} C)^2 (1 + \frac{\rho}{4} D)} + \frac{\frac{\rho}{4} (2(Q_{xk} - Q_{uv}) + 2t_2 (S_x - S_u)) s(S_x, Q_{xk}, S_u, Q_{uv})}{(1 + \frac{\rho}{4} C) (1 + \frac{\rho}{4} D)^2} \quad (78)$$

$$P(\mathbf{C} \rightarrow \mathbf{E} | X, Z) = \left( \det \left( \mathbf{I}_2 + \frac{\rho}{4} |Z|^2 \mathbf{R}_{t,2} \tilde{\mathbf{E}} \right) \right)^{-1} \left( \det \left( \mathbf{I}_2 + \frac{\rho}{4} |X|^2 \mathbf{R}_{t,1} \tilde{\mathbf{E}} \right) \right)^{-1} \\ = \prod_{k=1,2} \left( 1 + a_{k,2} |Z|^2 \right)^{-1} \left( 1 + a_{k,1} |X|^2 \right)^{-1}, \quad (82)$$

## REFERENCES

- [1] B. Clerckx and C. Oestges, "MIMO Wireless Networks: Channels, Techniques and Standards for Multi-Antenna, Multi-User and Multi-Cell Systems," Academic Press (Elsevier), Oxford, UK, Jan 2013.
- [2] D. Gesbert, M. Kountouris, R.W. Heath, C.-B. Chae, T. Salzer, "Shifting the MIMO paradigm," IEEE Signal Processing Magazine, vol. 24, no. 5, pp. 36-46, Sept 2007.
- [3] M. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," IEEE Trans. Inf. Theory, vol. 58, no. 7, pp. 4418-4431, 2012.
- [4] M. Kobayashi, S. Yang, D. Gesbert, X. Yi, "On the Degrees of Freedom of time correlated MISO broadcast channel with delayed CSIT," IEEE International Symposium on Information Theory, 2012.
- [5] S. Yang, M. Kobayashi, D. Gesbert, and X. Yi, "Degrees of freedom of time correlated MISO broadcast channel with delayed CSIT," IEEE Trans. Inf. Theory, vol. 59, no. 1, pp. 315-328, 2013.
- [6] T. Gou and S. Jafar, "Optimal use of current and outdated channel state information: Degrees of freedom of the MISO BC with mixed CSIT," IEEE Comm. Letters, vol. 16, no. 7, pp. 1084-1087, July 2012.
- [7] R. Tandon, S. Jafar, S. Shamai Shitz, and H. Poor, "On the synergistic benefits of alternating CSIT for the MISO broadcast channel," IEEE Trans. Inf. Theory, vol. 59, no. 7, pp. 4106-4128, 2013.
- [8] J. Chen and P. Elia, "Degrees-of-freedom region of the MISO broadcast channel with general mixed-CSIT," vol. arxiv/1205.3474v1, May 2012.
- [9] J. Chen and P. Elia, "Can imperfect delayed CSIT be as useful as perfect delayed CSIT? DoF analysis and constructions for the BC," in 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton), 2012, pp. 1254-1261.
- [10] X. Yi, S. Yang, D. Gesbert, M. Kobayashi, "The Degree of Freedom Region of Temporally Correlated MIMO Networks with Delayed CSIT," IEEE Trans. Inf. Theory, vol. 60, no. 1, pp. 494-514, Jan 2014.
- [11] X. Yi and D. Gesbert, "Precoding methods for the MISO broadcast channel with delayed CSIT," IEEE Transactions on Wireless Communications, vol. 12, no. 5, pp. 1-11, 2013.
- [12] J. Wang, M. Matthaiou, S. Jin, and X. Gao, "Precoder design for multiuser MISO systems exploiting statistical and outdated CSIT," IEEE Trans. on Comm., vol. 61, no. 11, Nov. 2013.
- [13] M. Dai and B. Clerckx, "Transmit Beamforming for MISO Broadcast Channels with Statistical and Delayed CSIT," submitted for publication.
- [14] J. Chen, R. Knopp, and P. Elia, "Interference alignment for achieving both full DoF and full diversity in the broadcast channel with delayed CSIT," in 2012 IEEE International Symposium on Information Theory Proceedings (ISIT), pp. 1887-1891.
- [15] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," IEEE Trans. Inf. Theory, vol. 45, no. 7, pp. 1456-1467, Jul. 1999.
- [16] B. Clerckx, G. Kim, and S.J. Kim, "Correlated Fading in Broadcast MIMO Channels: Curse or Blessing?," IEEE Globecom 2008, Dec 2008.
- [17] B. Clerckx, L. Vandendorpe, D. Vanhoenacker-Janvier, and C. Oestges, "Optimization of non-linear signal constellations for real-world MIMO channels," IEEE Trans. Signal Processing, vol. 52, no. 4, pp. 894-902, April 2004.
- [18] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," IEEE Trans. Inf. Theory, vol. 44, no. 2, pp. 744-765, Mar. 1998.
- [19] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," IEEE Trans. Inf. Theory, vol. 49, no. 5, pp. 1073-1096, May 2003.
- [20] H. Yao, "Efficient Signal, Code, and Receiver Designs for MIMO Communication Systems," PhD thesis, MIT, 2003.
- [21] M. Abramovitz, I.A. Stegun, "Handbook of mathematical functions with formulas, graphs, and mathematical tables," New York, Dover Publications, 1972.
- [22] S. Tavildar and P. Viswanath, "Approximately universal codes over slow fading channels," IEEE Trans. Inf. Theory, vol. 52, no. 7, pp. 3233-3258, Jul. 2006.
- [23] T.M. Cover and J.A. Thomas, "Elements of Information Theory," second edition. New York: Wiley-Interscience, 2006.
- [24] B. Clerckx, C. Oestges, L. Vandendorpe, D. Vanhoenacker-Janvier, and A.J. Paulraj, "Design and Performance of Space-Time Codes for Spatially Correlated MIMO Channels," IEEE Trans. on Comm., vol. 55, No. 1, pp. 64-68, Jan. 2007.
- [25] B. Clerckx and C. Oestges, "Finite-SNR Performance Analysis of Space-Time Coding in Correlated Ricean MIMO Channels," IEEE Trans. on Info. Theory, vol. 53, No. 12, pp. 4761-4777, Dec 2007.
- [26] B. Clerckx and C. Oestges, "Space-Time Code Design for Correlated Ricean MIMO Channels at Finite SNR," IEEE Trans. on Sig. Proc. vol. 56, No. 9, pp. 4365-4376, Sept 2008.
- [27] I.S. Gradshteyn, I.M. Ryzhik, "Table of Integrals, Series, and Products," seventh edition, Academic Press, 2007.
- [28] P. Dayal and M. K. Varanasi, "An optimal two transmit antenna space-time code and its stacked extensions," IEEE Trans. Inf. Theory, vol. 51, no. 12, pp. 4348-4355, Dec. 2005.